PHILOSOPHICAL REDIRECTION OF EDUCATIONAL RESEARCH

NSSE

NATIONAL SOCIETY
FOR THE STUDY OF EDUCATION

PHILOSOPHICAL REDIRECTION OF EDUCATIONAL RESEARCH

BUREAU OF EDUCATIONAL AND PSYCHOLOGICAL RESEARCH
GOVT. OF W. BENGAL
LIBRARY
CALCUTTA-19

Officers of the Society

1971-72

(Term of office expires March 1 of the year indicated.)

JEANNE CHALL
(1974)
Harvard University, Cambridge, Massachusetts

N. L. GAGE
(1972)
Stanford University, Stanford, California

JOHN I. GOODLAD
(1973)
University of California, Los Angeles, California

ROBERT J. HAVIGHURST
(1974)
University of Chicago, Chicago, Illinois

HERMAN G. RICHEY
(Ex-officio)
University of Chicago, Chicago, Illinois

HAROLD G. SHANE
(1972)
Indiana University, Bloomington, Indiana

RALPH W. TYLER
(1973)
Director Emeritus, Center for Advanced Study in the Behavioral Sciences
Stanford, California

Secretary-Treasurer

HERMAN G. RICHEY
5835 Kimbark Avenue, Chicago, Illinois 60637

C

PHILOSOPHICAL REDIRECTION
OF
EDUCATIONAL RESEARCH

The Seventy-first Yearbook of the National Society for the Study of Education

PART I

By
THE YEARBOOK COMMITTEE
and
ASSOCIATED CONTRIBUTORS

S.C.E.R.T., West Bengal
Library
Calcutta
25/3, B. C. Road

Edited by
LAWRENCE G. THOMAS

Editor for the Society
HERMAN G. RICHEY

BUREAU OF EDUCATIONAL AND PSYCHOLOGICAL RESEARCH
GOVT. OF WEST BENGAL
LIBRARY
CALCUTTA-19

19 NSSE 72

Distributed by THE UNIVERSITY OF CHICAGO PRESS • CHICAGO, ILLINOIS

4·2·75
2434

The responsibilities of the Board of Directors of the National Society for the Study of Education in the case of yearbooks prepared by the Society's committees are (1) to select the subjects to be investigated, (2) to appoint committees calculated in their personnel to insure consideration of all significant points of view, (3) to provide appropriate subsidies for necessary expenses, (4) to publish and distribute the committees' reports, and (5) to arrange for their discussion at the annual meeting.

The responsibility of the Society's editor is to prepare the submitted manuscripts for publication in accordance with the principles and regulations approved by the Board of Directors.

Neither the Board of Directors, nor the Society's editor, nor the Society is responsible for the conclusions reached or the opinions expressed by the Society's yearbook committees.

370.78
NAT

Library of Congress Catalog Number: 6-16938

Published 1972 by
THE NATIONAL SOCIETY FOR THE STUDY OF EDUCATION

5835 Kimbark Avenue, Chicago, Illinois 60637

Copyright, 1972, by HERMAN G. RICHEY, Secretary

The National Society for the Study of Education

No part of this Yearbook may be reproduced in any form without written permission from the Secretary of the Society

E R.T., West Bengal
4. 2. 75.
No. 2434

First Printing, 10,000 Copies

Printed in the United States of America

The Society's Committee on Philosophical Redirection of Educational Research

HAROLD B. DUNKEL
Professor of Education
University of Chicago
Chicago, Illinois

D. BOB GOWIN
Professor of Education
Cornell University
Ithaca, New York

LAWRENCE G. THOMAS
(Chairman)
Professor of Education
Stanford University
Stanford, California

Associated Contributors

HARRY S. BROUDY
Professor of Philosophy of Education
University of Illinois
Urbana, Illinois

DAVID W. ECKER
Professor of Art Education
New York University
New York, New York

DAVID HAWKINS
Professor of Philosophy
University of Colorado
Boulder, Colorado

BRIAN HOLMES
Reader in Comparative Education
University of London
London, England

EUGENE F. KAELIN
Professor and Chairman, Department of Philosophy
Florida State University
Tallahassee, Florida

LEONARD I. KRIMERMAN
Associate Professor of Philosophy
University of Connecticut, Inner College
Storrs, Connecticut

JAMES E. MCCLELLAN
Professor of Education
Temple University
Philadelphia, Pennsylvania

HUGH G. PETRIE
Associate Professor of Philosophy of Education
University of Illinois
Urbana, Illinois

MICHAEL SCRIVEN
Professor of Philosophy
University of California
Berkeley, California

IRA S. STEINBERG
Associate Professor
Oberlin College
Oberlin, Ohio

KENNETH A. STRIKE
Assistant Professor of Education
Cornell University
Ithaca, New York

Table of Contents

PAGE

OFFICERS OF THE SOCIETY, 1971-72 ii

THE SOCIETY'S COMMITTEE ON THE PHILOSOPHY OF EDUCATION . . v

ASSOCIATED CONTRIBUTORS v

INTRODUCTION, *Harold B. Dunkel, D. Bob Gowin, Lawrence G. Thomas* 1

SECTION I

Methodological Questions

CHAPTER

I. IS EDUCATIONAL RESEARCH DISTINCTIVE? *D. Bob Gowin* . 9

II. EXPLAINING AND UNDERSTANDING: THE IMPACT OF SCIENCE ON OUR CONCEPT OF MAN, *Kenneth A. Strike* 26

III. THEORIES ARE TESTED BY OBSERVING THE FACTS: OR ARE THEY? *Hugh G. Petrie* 47

SECTION II

Criticizing Current Concepts

IV. WANTED: NEW PARADIGMS AND A NORMATIVE BASE FOR RESEARCH, *Harold B. Dunkel* 77

V. OBJECTIVITY AND SUBJECTIVITY IN EDUCATIONAL RESEARCH, *Michael Scriven* 94

Part I: Objectivity and Subjectivity: The Foundations 94

Part II: Objectivity-Subjectivity: Methodological Applications. 116

VI. BEHAVIORAL DEFINITION OF EDUCATIONAL OBJECTIVES, *Ira S. Steinberg* 143

VII. INDIVIDUALIZED INSTRUCTION: A PROJECTION, *James E. McClellan* 164

VIII. CONCEPTIONS OF CULTURE AND SOCIETY IN EDUCATIONAL RESEARCH, *Brian Holmes* 193

SECTION III

Developing New Models

CHAPTER PAGE

IX. THE LIFE USES OF SCHOOLING AS A FIELD FOR RESEARCH, *Harry S. Broudy* 219

X. A MODEL FOR MAKING AND TESTING VALUE JUDGMENTS, *Lawrence G. Thomas* 239

XI. THE LIMITS OF AESTHETIC INQUIRY: A GUIDE TO EDUCATIONAL RESEARCH, *David W. Ecker* and *Eugene F. Kaelin* . . . 258

XII. HUMAN NATURE AND THE SCOPE OF EDUCATION, *David Hawkins* 287

XIII. AUTONOMY: A NEW PARADIGM FOR RESEARCH, *Leonard I. Krimerman* 327

INDEX 357

INFORMATION CONCERNING THE SOCIETY 367

LIST OF PUBLICATIONS OF THE SOCIETY 369

Introduction

HAROLD B. DUNKEL, D. BOB GOWIN, AND LAWRENCE G. THOMAS

Philosophy of education has been the theme of two previous yearbooks of this Society, the forty-first and the fifty-fourth. In both of those yearbooks, each chapter took a systematic philosophical position and then recommended distinctive ways of carrying on the educational process in the light of that position. The resulting marriage of philosophy and policies for practice became a theory to teach by, to administer by, to counsel by. Such theories were the main contribution of philosophy to the educational enterprise in the middle third of this century.

In recent years educational philosophers have been turning to quite different problems and hence offering a quite different contribution to educational thought. Instead of building prescriptive theories, they have been analyzing the meanings of educational concepts. Instead of deducing connections between conceptual systems and educational programs, they have been locating and examining the assumptions, the presuppositions, and the ambiguous beliefs that have long underlain educational policies and practices. Instead of inventing metaphysical systems and untestable explanations, they have been developing new models and paradigms to guide educational practice, including the important practice of educational research.

It is still too soon to estimate the effect of this analytical movement in philosophy on such familiar educational activities as classroom techniques and administrative procedures. We suspect the impact has not yet been very significant, although recent studies by philosophers of the concepts of equality and of equal opportunity in education have a great potential for clarifying certain aspects of school policy in administration and curriculum.[1] Perhaps in another year a yearbook can be devoted to this kind of influence. However, the effect of conceptual analysis on the definition of research problems underlying the educational enterprise has been considerable and promises to become even more important.

The conception of teaching, for example, has received a remarkably thorough analysis in the last dozen years at the hands of educational philosophers.[2] At the same time, the character of empirical research into the processes and effectiveness of teaching has changed radically, apparently following the lead of the philosophical analysts. In place of earlier attempts to establish factual connections between some global conception of "good teaching" and such other factors as courses taken, grades earned, and IQ possessed, recent research has concentrated on teaching styles, kinds of teaching acts, and the constituitive elements of each. Moreover, it is not at all uncommon to find both philosophers and empirical researchers contributing reciprocally challenging essays to a book of advanced readings on the nature of teaching.[3] The only other educational topic which comes close to receiving this much attention from philosophers as well as empirical researchers in recent years is curriculum, although so far their respective insights and concerns seldom appear as joint efforts.[4]

---

[1] E.g., an article by Paul Komisar and Jerold R. Coombs, "The Concept of Equality in Education," (*Studies in Philosophy and Education* 3 [1964]: 223-44) provoked a lively response of some half-dozen articles in subsequent issues of that journal, which as a whole constitute a very thorough analysis of this major educational principle.

[2] Beginning with B. O. Smith and Milton Meux, *A Study of the Logic of Teaching*, Bureau of Educational Research, University of Illinois (undated but about 1958) and running through Thomas F. Green, *The Activities of Teaching* (New York: McGraw-Hill Book Co., 1971).

[3] Nathaniel Gage, ed., *Handbook of Research on Teaching*, AERA, (Chicago: Rand McNally & Co.), 1963.
Ronald T. Hyman, *Contemporary Thought on Teaching*, (Englewood Cliffs, N.J.: Prentice-Hall, 1971).

[4] For leading examples by philosophers, see R. S. Guttchen and B. Bandman (eds.), *Philosophical Essays on Curriculum* (Philadelphia: J. B. Lippincott Co., 1969) and Jane R. Martin (ed.), *Readings in the Philosophy of Education: A Study of Curriculum* (Boston: Allyn & Bacon, 1970). For good examples of curriculum building based on research and analysis of practice, see Edward Krug, *The Secondary School Curriculum* (New York: Harper & Bros., 1960) and Hilda Taba, *Curriculum Development: Theory and Practice*, (New York: Harcourt, Brace & World, 1962).

The present volume is written by philosophers, most of whom are already working closely with empirical researchers. But the topics they are undertaking to analyze here have not been extensively explored before. Their essays, accordingly, are invitations to other empirical researchers in the reading audience to respond with their doubts, their questions, and their alternatives.

The redirection of educational research requires more than cooperation with current empirical researchers. It also calls for the identification of significant educational problems through the interaction of philosophers with teachers and school administrators. Since these latter persons make up the majority of the professional membership of NSSE, these chapters are also addressed to them with an invitation to participate in framing significant questions for inquiry. Do these chapters clarify what may legitimately be expected from empirical research in education? Do they open up conceptual areas where empirical research, in the judgment of professional workers in education, is especially needed?

In spite of the diversity of topics among the chapters which follow, one criterion was a common demand for each chapter: How can this analysis improve the practice of educational research? The chapters to follow could be arranged in several different orders. We have chosen one that highlights the three distinctive ways in which philosophical research can contribute to the improvement of empirical research in education.

The chapters in Section I attempt the improvement by raising broad questions of methodology. Gowin's opening chapter proposes a conception of educational research that would make this realm of investigation more normative, more significant, and more demanding. Strike's chapter develops the methodological difference between the scientific explanation of human behavior and the humanistic understanding of human actions. Petrie's chapter reveals the impossibility of observing any kind of facts without some conceptual set which screens and interprets the observations.

The chapters in Section II attempt to meet the criterion of improving educational research by criticizing certain conceptions and research practices now in common use. Dunkel's chapter discloses the inadequacies of the traditional physical science paradigm for defining significant problems to investigate for the improvement of education. Scriven's chapter shows how two distinct meanings of objectivity became confused with each other and distorted the entire process of finding rational explanations of human behavior. Steinberg's chapter examines the logical grounds for behavioral definitions of educational goals and shows how present uses of this technique have become both faulty and dangerous. McClellan's chapter identifies the many, sometimes conflicting, meanings of individualized instruction and clarifies the significant choices for the researcher. Holmes in his chapter undertakes to show through two major examples of cross-cultural research how a sharper preliminary analysis of such concepts as "society" and "culture" would have increased the significance of these two studies.

The chapters in Section III seek to improve educational research by developing new or neglected models that deserve more attention in this kind of empirical investigation. Broudy's chapter directs attention to two familiar uses of school achievement that are comparatively neglected in present forms of achievement testing. Thomas, in his chapter, argues that all value claims can be converted into empirical questions and proposes a model for the empirical study of value judgments. The chapter by Ecker and Kaelin provides a model for organizing and classifying the empirical research that can be done in aesthetics. Hawkins, in his chapter, takes three dimensions of human nature—equality, freedom, and rationality—and reconceptualizes them to guide distinctively different empirical investigations. Krimerman's chapter explores the nature of human autonomy and defines several fresh parameters needing empirical exploration.

Before closing these remarks, we wish to express our special thanks to Herman G. Richey, Secretary-Treasurer of the NSSE, who readily took our organizational problems to the sympathetic ears of the Board of Directors and who gave us steady encouragement and prompt financial aid during the several years of planning this volume. Our thanks also to Nathaniel L. Gage, the member of the NSSE Board who carried the specific liaison responsibilities between our editorial committee and the Board, succeeding admirably in keeping both groups happy. We also express our appreciation to the Department of Education at Cornell University for assisting in the planning and support of a conference of authors held there

September 1-3, 1970, at which time early drafts of most of these papers were read and criticized. Thanks are also owed to the USOE Title IV Research Training Program at Cornell, which provided some of the funds for this conference, and to the dozen graduate fellows who participated in the deliberations, read all the papers, and offered many valuable suggestions to the authors and the editors. We feel that these efforts helped greatly in creating a commonly understood purpose for the yearbook.

The improvement of all the various activities that make up educational practice has been for years a hope of both empirical researchers and educational philosophers. The editors of this yearbook hope that these chapters will lead to conjoint efforts among philosophers, researchers, and educators which will have synergistic effects on the improvement of all educational practices. Yet we accept the pragmatic test, which resides with the readers, as to whether this volume meets their expectations and fulfills our intentions.

SECTION I

METHODOLOGICAL QUESTIONS

CHAPTER I

Is Educational Research Distinctive?

D. BOB GOWIN

Any day that some agency grants money for educational research is a day someone thinks he knows which proposals qualify as research proposals and which do not qualify. Every doctoral dissertation in education, every published article in a refereed educational journal, every attempt to improve educational practice through disciplined inquiry represents the use of criteria to judge research processes and products as belonging in the domain of education. Where do these criteria come from and what are good and bad consequences of their use?

Most commonly, criteria for creating, directing, and judging educational research come from concepts of science (philosophy of science) and the customs of research practice. Usually the researcher tries to get clear about the procedures of research practice before undertaking to use them in an educational setting. The main recommendation of this paper is that this familiar pattern be changed. The researcher should first try to be clear about the concepts, methods, and procedures of educational practice so as to be able to select phenomena to study that pass as educational phenomena and then adapt, invent, or utilize relevant research procedures. The reason for this recommendation is simply that many events which are educational never get studied now, and many events which educational researchers concern themselves with now have little or nothing to do with education. Further, to follow this recommendation would force researchers to argue first about what is and is not educational, rather than discussing only what is and is not scientific. This kind of discussion would lead to a thorough analysis of educational theories, concepts, and practices.

It is my belief that this analysis would reveal a most significant fact about *educational* phenomena: They are man-made (artifactual), not natural. They are therefore not likely to yield laws and other modes of invariance such as the natural sciences report in that domain. Whatever regularities researchers are to find in educational phenomena will have been determined by human beings in a social context. Normative judgments (rules, policies, value judgments, ideals which govern action) condition greatly the phenomena to be studied. Change a belief system and the content of research reports will be very different.

The entire first chapter of Travers's newly revised text[1] is based on the uncriticized assumption that educational research is to be like research in natural sciences—to search for generalizations (laws) which are expected to hold in a variety of situations. Travers designates as relatively poor the research done on the reading of textbooks because it will no longer be valuable when textbooks vanish from the educational scene. It would have been better, Travers claims, if the researchers had dealt with learning from all kinds of documents because these generalizations would last longer or apply to more situations. But any educational research generalization will become largely worthless when men change educational patterns.

Take, for example, the two-thirds "rule" which states that two-thirds of the time in a classroom someone is talking, and during two-thirds of that time it is the teacher that is doing so. This regularity in educational phenomena has been consistently observed and recorded over decades. It looks like an empirical generalization that will hold up like generalizations in chemistry or some other science. But change the educational concept from self-contained classrooms to open learning centers, from group instruction to personalized instruction—change the educational concepts which direct the educational practice and the regularity expressed in the two-thirds rule will vanish from the educational scene.

The fact that educational phenomena are man-made events rather than events "given in nature" has a number of consequences for the distinctiveness of educational research. One consequence is that educational research will be unusually diverse. Consider the abbreviated examples of research which follow and try to determine what criteria are being used to judge that these examples qualify as educational research.

---

[1] Robert M. W. Travers, *An Introduction to Educational Research*, 3d. ed. (New York: Macmillan Co., 1969), chap. 1.

An empirical researcher computes a +.30 correlation between sex and political opinion—males more liberal and females more conservative—in a group of 1200 students from twelve colleges selected as being representative of institutions of higher education in America.

A professor of comparative government writes a book on Southeast Asia and devotes a short chapter to descriptive information from various countries on the number of students in school, the ratio of teachers to students, the average level of achievement of teachers, and the estimates of national expenditures for education.

A psychologist at a leading university devises a fifty-item teacher rating scale (e.g., "Has a sense of humor"; "Lectures are over my head, about right, too elementary") and half the faculty members agree to use it; the results are compiled and computer print-outs are made available (anonymously to the public; personally to the individual instructor).

A sociologist visits a boys' camp in the summer and devises with the boys a "game" which pits two groups against each other. One group succeeds against great obstacles and the other group fails even under favoring conditions. The researcher publishes a study on contrasting leadership styles.

A philosopher writes a book on logic in which all the examples involve teachers and acts of teaching. Another writes a carefully thought-out argument distinguishing the several meanings of "know," "knowing," the "known," and the "knower." Still another philosopher writes a book on semantics including symbol systems in art products (musical scores, choreographic charts, etc.) and ends his book with a paragraph on the significance of a concept of symbol systems for education.

A political activist and journalist visits a newly established free school and keeps a journal of events as he sees them over a period of a year. He publishes his observations in newspapers and magazines and receives an award for excellence in educational writing.

A veteran poet describes how he taught elementary pupils in an inner-city school to be poets. A beginning assignment asked each pupil to write a single line, following the agreed-upon rules that each line should contain a color word, a reference to a comic strip character, and the name of a city or country, and that the line should begin with the words, "I wish." The teacher collected the lines, shuffled them, and read them back to the class as a poem. He reported that the children were excited, were talking, waving, blushing, laughing, and bouncing up and down. He concluded that the experience was a good beginning experience because it made the pupils feel like poets, and it made them want to write more. The poet's essay continues with reports of things he did, with examples of pupils' poems, with his judgments about the poems, and with observations on his pedagogical practices.

A black leader writes his autobiography telling of his change from a petty thief, pimp, and drug addict to a person in a position of political power. He notes the reading he did in jail, the techniques he used to "kick" the drug habit, and his recommendations for a new life-style.

These examples could have real names, places, and dates attached to them. By using some criteria rather than others each example would qualify as educational research. At the same time we note that the use of some criteria rather than others provides us with a basis for criticism of each research product. In the analysis which follows note that criteria come from both educational and research domains.

In the first example we are faced with the point that for some research products the mere fact that the work was done in the setting of an educational institution is sufficient to classify it as educational research. (But note that we would not readily classify as medical research the correlational study between sex and political opinion if it was done in hospitals.)

The demographic and fiscal data about students, teachers, and budgets is easily quantifiable and gives the appearance of direct comparison from one country to another; but, without some further idea about what happens when students and teachers actually interact in the educational setting, we can make no valid educational inferences about the educational processes in the different countries. (The French concept of education is very different from the British concept, although this way of characterizing educational concepts needs much qualification.)

Closer to the educational scene are the familiar rating scales of college teachers. We get some information from these studies that is potentially useful to the individual instructor as he tries to improve his teaching, but the summative form of evaluation of instruction leaves us in the dark about the events, episodes, and processes that occurred *during* the term of instruction (formative evaluation). Ratings taken at the end of a semester usually fail to take account of the process of instruction during the semester. Lectures over the heads of students at the beginning might have become too elementary by the end.

The educational sociologist who uses an informal setting, like a summer camp, to rig experimental events can come close to finding those variables in group processes which might be manipulated for educational purposes in a formal setting. One very important educational concern—the teaching of subject matter—is missing.

In the case of the book on logic, using examples involving teachers and acts of teaching, most empirical researchers would not think of it as research at all. But, granted that we could call it conceptual research, still some identifiable concept of teaching is needed. If teaching is conceived of as guidance of human conduct in ethical, political, and social matters, then a logical analysis of the advantages and disadvantages of the practical syllogism would be a contribution to the study of teaching. If teaching is held to be a careful explanation of key concepts in a subject like physics, then logical analysis of scientific explanation would be called for. But simply finding illustrations of principles of logic in the behavior of teachers does not make the work a contribution to educational research; we need to know more about the logical problems found in the subject matter that teachers use in their acts of teaching, and this calls for a specific concept of teaching—the teacher's view of subject matter.

While there is face validity to the claim that education concerns knowing, the knower, and the known, there is no necessary reason why the distinctions the philosophers make about these epistemological predicates carry any educational significance. One would have to see if the distinctions made in the philosophical context actually did carry over into the educational context. This comparing of intellectual and practical contexts requires a special study of its own; the same thing may be said for the semantics of symbol systems.

The educational journalist is helpful as a researcher to the extent that he records events in the educational context, but we cannot tell whether his political ideology or his educational concepts govern the events to be selected and reported. The presence of two or three observers would lead toward greater intersubjective agreement, but any events which are newsworthy more than educational would color the report considerably.

The veteran poet demonstrates his skill and illustrates his poetic and critical judgments through the medium of the children's own work. Other competent judges of poetry can use their standards of literary evaluation on the pupils' "artistic" works and can conclude independently of the poet-teacher whether the educational events he brought about by his teaching are worthy or not. So it should be possible to conclude from his report that educational events are occurring. We are restricted from making very broad empirical judgments, however, since the veteran poet may be an exceptionally talented and idiosyncratic teacher, doing things that most teachers could not do. But we are getting closer still to the actual scene of a teacher, a subject, pupils, and the involved triadic relationship that develops during an educational episode.

Finally, the autobiography of an exceptional person engaging in the complex and difficult task of self-education—where teacher, pupil, and the selection of subject matter are all wrapped up in one person—scarcely seems a topic for empirical educational research. But the record can be checked with the techniques of historical and case studies and, insofar as the individual is like many other individuals, the procedures he records of dramatic self-education might well be used in a large number of other settings by other individuals. These further developments could be empirically observed and recorded.

We can see in these truncated examples a wide range of studies and sources of understanding. There is a range of studies from those involving large-scale and group data to single studies of single individuals. There is a range from intellectual to practical contexts, from self-reports to journalistic reporting. There is a range of criticisms using different criteria of evaluation.

By working through these examples we have amassed *a cluster of criteria*. We have the beginnings of distinctiveness. To determine if a piece of research qualifies as educational research, look for criteria derived from the following:

1. *The social setting.* Is it an educational institution? Is it a rule-governed social setting conducive to educational processes?
2. *The concept of education.* Different countries, social orders, social groups have different conceptions of education. Determine what concept (or concepts) of schooling and education appears to be working in the given social setting.
3. *The educative process.* Studies which report data only on the

products of the educative process are likely to mask significant events in the educative episodes. The educative process is almost always a process of "from . . . to." That is, the process is developmental, episodic, continuous. A key feature of an individual being educated is that the person begins in some state of immaturity, lack of skill, ignorance, uninformed attitude, and so forth, and moves through the process in a variety of ways toward some ending point of maturity, skill, knowledge, changed belief, and so on. Determine what unit of time is appropriate to the educative process.

4. *Content or the subject matter of instruction.* In many cases where psychologists or sociologists manipulate psychosocial variables there is little or no concern given to the role of substantive knowledge in the events which are studied. Determine what counts as instructional content.
5. *The context of inquiry.* Studies in logic or epistemology or in any other field outside of education may be valid in themselves, but, unless a special study is made bringing findings over into the educational context, nothing follows for educational practice.
6. *The person.* An instance of dramatic education, such as Helen Keller or Malcolm X, makes us realize that the human agent is always involved if education is taking place. Their victories over incomprehension may only represent more dramatic instances of what every child undergoes in the process of education. In any event it takes a person who has understood something (pain, how to read) to recognize that another person understands it also.

I have called these criteria a cluster. I do not wish to claim that this cluster is an exhaustive list. (Teachers, for example, are not directly mentioned.) Nor do I wish to claim that a more complete cluster of criteria would represent the *essential* characters of educational phenomena. In an essay on scientific method and the nature of science, Black argues against trying to find the essential character of science. He says that the difficulty lies in the traditional (Aristotelian) view of definition which assumes everything has a determinate essence and hence that there is only one definition appropriate to it. Black writes:

> Neither observation, nor generalization, nor the hypothetico-deductive use of assumptions, nor measurement, nor the use of instruments, nor mathematical construction, nor all of them together . . . can be regarded as essential to science. . . . Astronomy makes no experiments, mathematics uses no observation, geography is mainly descriptive (not many generalizations), much classification and taxonomy frames

no abstract generalizations, biology is just beginning to use mathematical idealization and formalization. The characters mentioned are neither necessary nor sufficient. They are present or absent in degree. Their joint presence in high degree creates conditions recognized as preeminently scientific.[2]

Something of the same thing is what I am suggesting for determining distinctively educational phenomena. We might say that the joint presence in high degree of elements of a rule-governed social setting, a clear concept of education, a humane process for transforming the uneducated into better educated individuals, bodies of knowledge as subject matter, and so on (I leave it open deliberately), create conditions recognized as education.

We still have the problem of educational research. I do wish to suggest that the procedure I have used, albeit sketchily, is a valid way to develop a theory of educational research. A long time ago Dewey suggested that the theory of inquiry should be derived from a disciplined reflection upon the successes and failures of actual pieces of research. Schwab has done a work of this kind in "What Do Scientists Do?"[3] This recommended procedure, when well done, is in contrast to the usual procedure which derives grounds for acceptable scientific research from the philosophy of science.

We thus come to our first conclusion: Educational research is distinctive when criteria from *both* education and research can be jointly used. Why should such a tediously obvious thing need to be said? The reason for saying it is that many writers on the topic of educational research do not say it. Worse still, what they do say about educational research quite often leads us away from the search for educational criteria. Suppose someone said he was doing biological research. The label "biological research" leads us to expect research dealing with living things because of our understanding of what "biological" means. But if one said that he was doing behavioral research, we would not necessarily think it concerned education.

Writing in the *Encyclopedia of Educational Research* on the topic "Research in Education," Professor Kerlinger takes the posi-

---

2. Max Black, *Problems of Analysis* (Ithaca, N.Y.: Cornell University Press, 1954), p. 13.

3. Joseph J. Schwab, "What Do Scientists Do?", *Behavioral Science* 5 (January 1960): 1-27.

tion that educational research is social scientific and behavioral research applied to educational problems. Why is it this? "It is social scientific research for a simple reason: an overwhelming majority of its variables are psychological, sociological or social-psychological" (p. 1127).[4] Examples of these variables are achievement, aptitude, motivation, intelligence, reinforcement, social class, and race. But can we say what is distinctive about educational research by identifying only such *variables?*

Suppose we read scientific reports about drugs and learning. One variable is the drug; the other is the speeding up or slowing down of learning rates. Is this research therefore biochemical research, educational research, or both? We have a long list of excellent research studies on the growth and development of children which concern variables we would classify as physiological, such as maturation rates of boys and girls. This information is useful in a number of educational contexts. Since the rate of growth is a variable, then are we limited to classifying the research as physiological research? Or can it be educational research as well? We cannot tell by using Kerlinger's criterion.

Another variable which has a powerful effect on learning in schools is the way the space is designed. The research work architects do in designing space may or may not take into account educational problems and so it would be impossible to classify the research, knowing only what variables are involved. I conclude from these counterexamples that Kerlinger's reasoning is in error, that we cannot decide that educational research is social scientific and behavioral because *some* of its variables can be described in this way. Much more than knowledge of the nature of variables is needed before we can identify distinctively educational research.

Worse still, Kerlinger's classification of educational research as social scientific and behavioral research leads us away from considering events which are clear cases of educational phenomena. A prime example very often overlooked by psychologists and sociologists is that of subject matter. A careful study of what qualifies as significant subject matter that ought to be taught to children can take a long time, involve many observations, generate many conceptual analyses and definitions, require much field testing and correction, involve large numbers of people in evaluation, and so forth. Why should curriculum studies not be counted as educational research? Another example concerns the concept of teaching. There are a number of competing concepts of teaching and the analysis of these concepts clearly deals with an educational problem, but the research deals neither with the social nor the behavioral because it deals with concepts, not social relations nor behavior change. Thus, by tacking on the term *social science* or *behavioral* to research we can be led away from considering important educational phenomena.

---

[4] Robert L. Ebel, ed., *Encyclopedia of Educational Research*, 4th ed. (New York: Macmillan Co., 1969), pp. 1127-44.

Kerlinger's position leads us to consider the conceptual and methodological problems of psychological and sociological research. He shares with Travers (and many others) the assumption that research is to discover laws, to make general explanations of natural events. If we recognize that educational phenomena are not the same kind of "natural events" as those found in the natural sciences, then we have little or no reason to suppose that the theory of science which guides research practice in those areas where laws have been found will be the correct one for educational research. The fact that so few laws have been found through psychological and sociological research should lead us to have some doubts about this view of science as the basis for educational research.

We should recognize that it is difficult to define either science or education. That is, it is not possible to give a definition in the form of a statement of the necessary and sufficient conditions for the correct application of the term *science* or *education*. One major reason for this difficulty is that both science and education are human activities; they are inventions generated out of human experience for the purposes of better control and understanding of human experience. If we know anything for sure about human experience, it is that it varies, that it is inventive, creative, in process, subject to change. Whatever is to be known about these complex processes will change as time moves on. Definitions will be made and revised. Accordingly we need conceptions of both science and education which will permit us to revise in a responsible way our conceptions about these phenomena.

I take from the empiricist tradition in philosophy two assumptions. First is the assumption that human experience supplies the basis and ground for intelligence and insight into that experience. The second is the assumption that any assertion made should have support, i.e., some evidence to supply warrant and backing for that assertion.

We thus come to a second obvious conclusion: Studies of educational phenomena should be done so that assertions about those phenomena are related to evidence used to back up those assertions. Why, we should again ask, is so obvious a criterion for what is to count as educational research needed? The reason for saying the obvious is that statements of the aim of scientific research in education given by many writers on this topic fail to stress this point. Indeed, the aim that is often selected is so limited and guarded that most research actually done will fail to achieve the aim recommended. Further, establishing in a careful way the relation between assertions and evidence is much more difficult than it appears. It is a common failure of most actual pieces of research. Examples are forthcoming, but first the philosophical point. It seems to be clear today that any set of complex events can be validly described under an almost infinite number of conceptual systems. This plurality comes about because the relation between concepts and the objects and events to which they refer is forever open.

For example, examine the classic definition of measurement by Stevens as "the assignment of numerals to objects and events according to rules." [5] Both the *assignment* of numerals and the *rules* to be followed are arbitrary; they are open and not of a determinant necessity. The basic process of measurement then relates back to the empiricist criterion: Give some way to interpret the connection between an assertion about phenomena and the evidence (or data) which backs that assertion. If we suppose we have measured intelligence, or anxiety, or school atmosphere, and have as a consequence been able to say something meaningful about those concerns, then we are obliged as empirical researchers to connect the assertions and the evidence. Fulfilling this empiricist tradition is difficult, as some examples will show.

---

[5] S. S. Stevens, "Mathematics, Measurement, and Psychophysics," *Handbook of Experimental Psychology* (New York: John Wiley & Sons, 1951), pp. 1-49.

Consider the following advice given to educational researchers.

> The word *empirical* connotes a direct observation. An *empirical question*, then, is one for which there is possibly a direct observation to determine its answer. Nonempirical questions require a combination of direct observation and inferences about the observation and related concepts. A couple of quick and simple examples will solidify these statements. How many test items on the arithemetic section of the *California Achievement Test* will be answered correctly by a specific child? To answer that question, the test is administered, scored, and the items marked correctly are counted. The number correct *is* the observation.[6]

The problem here, of course, is that counting test items is a direct observation of what? A person counting? The "number correct" is not observed; it is an inference from (a) the validity of the key and (b) the reliability of the scorer, which is crucial since the quotation does not say administered by the inquirer. Counting test items is not a direct observation of a child's behavior working arithmetic problems, nor of a child making change in a store, nor counting marbles. It may be a *record* of a child taking an arithmetic test, but it might just as well be a record of a child obediently following orders, or responding whimsically because bored or fearfully because threatened.

> Consider the same situation but a question which cannot be labeled "empirical." Does a specific child understand arithmetic? The answer to that question *is not* directly observable. Understanding is a *mental* activity not visible or countable in any direct manner. An answer can be generated; however, it requires the selection of some directly observable activity as a logical indicator of understanding. In this case we probably would look at *both* the number of items answered correctly *and* the grade level of the child.[7]

How, we should ask, is looking at test items passed (say, fifteen) and the grade level of a child (say, grade 2) any different? Neither is a direct observation of the child in an instructional setting (testing is not teaching arithmetic). Both appear to be equally direct (but not really direct at all) and equally poor indicators of a child's understanding of arithmetic. A teacher could assay a child's arith-

---

[6] William J. Gephart and Bruce B. Bartos, "Profiling Instructional Package," Occasional Paper No. 7, Phi Delta Kappa Research Service Center, August 1969, pp. 4-5.

[7] Ibid., p. 5.

S.C.E.R.T., West Bengal
Date 4. 2. 75.....
Acc. No. 2434.........

metical understanding by finding a situation in which a *child's* purpose (selecting eleven children for a football team) required the use of arithmetical concepts to fulfill the purpose. Or, a child can simply express his understanding directly by saying something like, "I can count by nines," and then doing the counting correctly out loud. Also, surely there are problems of the construct validity of the test itself. Further quibbles: Understanding seems to be more a mental state than a mental activity. (Calculating would be a mental activity.) What is chosen as directly observable would seem to be the *empirical* (not logical) indicater of understanding. And so on.

While we can say that in a sense the relation between evidence and assertion is made clear (a child understands arithmetic if a number of test items are passed), the advice has the bad consequence of leading us away from, rather than into, the *educational* phenomena of interest. We can establish the relation between assertion and evidence, but not the relation between assertion-evidence and the human experience of teaching children arithmetic.

In a paper with the beguiling title "Reconstruction of Educational Research," Professor Shulman makes a keen effort to relate research methods to schooling situations. He judges (correctly, I think) that the research model of the single variable systematically isolated and manipulated while all else is held constant just will not do any longer for educational research.

> Most conceivable schooling situations will possess certain common characteristics: (1) They involve the attempt to modify or manipulate a setting (with or without teacher) to bring about desired changes in a learner; (2) They take place over relatively extended periods of time; (3) They involve the simultaneous input of multiple influences and the likely output of multiple consequences—some predicated, others not; and (4) They are characterized by variability of reaction to ostensibly common stimuli, that is, not all learners learn equally or react similarly to specific acts of teaching.[8]

Would any or all of the following situations be acceptable as "conceivable schooling situations"?

1. A religious order indoctrinating received dogma to novices

---

[8] Lee S. Shulman, "Reconstruction of Educational Research," *Review of Educational Research* 40 (1970): 371-96.

S.C.E.R.T., West Bengal
Library
370.78
NAT
2434
4.2.75

2. A military training program conditioning recruits to blind conformity to commands of authority
3. A politically radical minority group instilling pride in the minority condition and insisting on a sloganized rhetoric of revolution
4. A commercial management training program set in a context of vicious competition for profits and advancement of individuals by any means consistent with a profitable outcome

Take the first characteristic: ". . . the attempt to modify or manipulate a setting (with or without a teacher) to bring about desired changes in a learner." All four examples qualify. The weasel word is *desired*. By what criteria of educational value are we to distinguish desires for religious indoctrination, military obedience, political radicalization, or commercial profit from desires for other sorts of education?

What changes in a learner are desired? Desired by whom—the pupil or the parent or the teacher? Is there a difference between the desired and the desirable? Dewey frequently emphasized this point in value-making.

We may also ask, what is changed? Is it behavior modification or a deepening of understanding? Does it make any difference what intellectual matters are discussed in these schooling situations—e.g., algebra, Greek, political theory, or *War and Peace*? To make any headway on any research design, prior analysis of what counts as educational phenomena will have to be done.

To Shulman's credit he follows the paragraph I quoted with the following statement showing his concern for the congruent relation between phenomena of education and a research setting.

> What is described above is a highly complex and variegated activity, involving students, subject-matter and sources of instruction in an educational setting. Any research which purports to deal systematically with phenomena at this level of complexity must itself reflect an appropriate level of complexity. The ideal research setting, to be congruent with the description of the educational setting offered above, must be (1) *experimental*; (2) *longitudinal*; (3) *multivariate* at the level of both independent and dependent variables, and consistent with that, (4) *differential*, in that the interactions of the experimental programs with the students' entering, individual differences are treated not as error variance, but as data of major interest in the research.[9]

---

9. Ibid.

There is nothing in Shulman's view which really helps us discriminate the defensible educational setting from any of the others. The problem is an old one with educational researchers. Methods of research do not tell us what is educational nor what is worth doing. Many experts on research methodology claim that they are not even trying to decide about defensible educational phenomena. It seems to me, however, that selecting the phenomena of interest is one of the most crucial determinants of the eventual research product. I think this selection cannot be done without the conjoining of educational concepts with conceptions of methods of research.

What are the bases for distinguishing between that subject matter which is capable of indoctrination and that subject matter which by its very nature is not so amenable to such use? We need criteria for distinguishing. Would it make no difference to the educational researcher to study the development of an open religious attitude instead of religious indoctrination, or to study the means of creating military morale within a progressive and democratic society instead of the conditions of conformity and blind obedience? Would it not be more important to study the broader educational setting of political socialization than the effects of radical rhetoric upon minority groups? Should not the greater concern for the whole with respect to the benefits of an economic system be better to study than the limited devices whereby individuals are taught to value profits only? To each of the broader alternatives something like a body of knowledge (a discipline) belongs. The narrower ones belong in the domain of educational research only when they are studied in connection with the broader alternatives.

The reconstruction of educational research cannot come about until this variable, subject matter, is more widely debated and, hopefully, its meaning for research better understood. A recent symposium at Stanford University[10] revealed through its examples the complex interplay between subject matter (science, art, English, social studies . . .), a concept of teaching (as heuristic teaching), and the variety of ways empirical research can be directed toward these phenomena. The historical perspective reveals in many cases that what was once acceptable doctrine (acceptable as presenting true knowledge claims) changes into unacceptable opinion. Every subject matter has its knowledge claims and value grounds reconstituted. It is within the meaning of subject matter, as representing human efforts to comprehend truly and know validly significant areas of human experience, that we find criteria at any time for distinguishing between truth and falsehood, between opinion and belief, between unquestioned dogma and enlightened doctrine.

---

[10] Richard E. Snow, ed., *A Symposium on Heuristic Teaching*, Technical Report No. 18, Stanford Center for Research and Development in Teaching, December 1970.

Summary

To decide how distinctive educational research is, we must select criteria for deciding what is to count as education. Determining what the phenomena of research interest are comes from identifying the operating educational concept. The events to be studied are not "given events in nature" but are man-made. As a consequence, the diversity of the social settings and the rules which govern them will generate distinctive features of educational research. That is, the network of relations between characteristic phenomena of interest, the educational concepts, and the empirical methods of inquiry all determine what is distinctive about knowledge claims produced by the effort of educational research.

Eight examples of debatable cases of educational research were sketched. In judging these cases as examples of educational research, criteria from both education and research were needed. The procedure of examining actual cases generated a cluster of criteria. No claim of exhaustiveness was made. Since standard criteria of science are derived from either clear cases or abstract idealizations, these criteria do not help us decide borderline cases. Thus, continual examining of research products is recommended as the way to develop standards for the direction and control of inquiry.

Two assumptions of the empiricist tradition were affirmed. First, from a study of human experience we can derive criteria for directing and controlling that experience whether it concerns science, education, or both, in the form of educational research. Educational research is seen as distinctive when criteria from both education and research are conjointly used. The second assumption helps us see that occasions for inquiry require the backing of as-

sertions by evidence, but the closeness of the relation between concepts and events makes this procedure difficult, open, and arbitrary.

In short, educational research is seen as distinctive when knowledge claims are set in the context of education so that the relations between educative event, assertion, and evidence are clear.

CHAPTER II

Explaining and Understanding: The Impact of Science on Our Concept of Man

KENNETH A. STRIKE

Aristotle held that man was a rational animal. He inferred from this notion of the nature of man that the life of the mind, the contemplative life, is the best life for man and that one is only genuinely happy and self-actualized when he is living such a life.

Such notions are, of course, part of the stuff out of which educational policy is generated. Much in our conceptions of what we want in education and in our notions of appropriate ways to achieve what we want depends on what we think man is or what he can be. If we think man is potentially rational, we attempt to devise ways to enable him to actualize this potential, or, if we think that man is basically good, we feel that he can be trusted with a large measure of freedom.

We may also feel that such classical positions about man are a bit outdated. Man really is much more like a machine than we had previously supposed. Perhaps he is a complex biological machine, but, nonetheless, he is a physical thing with moving parts, and his behavior can be undestood like the behavior of any other physical thing. If we think this, we may conclude that man should be treated like a machine. Perhaps we in education will want to talk about "programming" or conditioning our students, or be very interested in educational technology, and perhaps we will conclude that really what education is about is how we can most efficiently get our students to exhibit those sorts of behavior which we find desirable. The point to note here, however, is not that certain views of man imply or suggest certain positions about educational policy, but simply that in fact there is a general connection between conceptions of what man is or can be and positions about educational policy.

Very often, social or behavioral scientists hold positions about human nature. What is most noteworthy about this otherwise unnoteworthy fact is that such views are often not held as a result of empirical research into behavior, but rather are presupposed by and exert a formative influence on the kind of research which is done and on the notion of what counts as a suitable explanation of human behavior. If this is the case, then it follows that certain features of educational policy may be decided not by the results of empirical research into human behavior, but by the commitments we make prior to such research. I believe that this is in fact the case; however, the point of these remarks should be clear enough by the end of the paper. Thus, I shall pass on to the main topic for discussion.

I want to ask what it means to understand human behavior. More specifically, I want to ask what it means to understand a human action. (This remark presupposes that some kinds of behavior are not actions. I shall characterize an action as an instance of behavior which is done intentionally. Thus, going to the store to buy a loaf of bread is an action. Having a muscle spasm, falling down the stairs, or having an epileptic seizure are not. Things which happen to me are not actions, only things which I do are such.) I shall address this question by focusing on the notion of understanding (*verstehen*) as exhibited in the writings of Max Weber.[1] I shall attempt to show that Weber's notion of *verstehen* can be defended against several types of objections which have been raised against it. More important, however, I shall suggest that Weber's notion involves an acceptable conception of human nature, one that enables us to treat man as an object of scientific study without requiring us to give up many features of a classical and humane view of man. Finally I shall try to suggest what this means for some broad issues of educational policy.

---

[1] My reasons for directing my attention to Weber are that much of the literature on the concept of *verstehen* has been concerned with his writings and that I think that many of the general features of his views are defensible. Thus, it seemed appropriate to direct myself toward the concept of *verstehen* by considering Weber's work and some of Weber's critics. Let me hasten to add, however, that my primary purpose is not Weber exegesis, but rather the development of an adequate view of *verstehen*. Weber's writings are most complex and often manifestly unclear. When exegetical difficulties arise, I shall accordingly provide that interpretation which I feel to be most defensible. Further, I shall not be overly careful to point out where Weber ends and Strike begins. Hopefully this approach will add to the brevity of the text. I trust the reader will find this sufficient justification.

The basic dispute clustering around the notion of *verstehen* has typically sounded something like the following: The advocate of some version of the *verstehen* doctrine will claim that human beings can be understood in a manner that other objects of study cannot. Men have purposes and emotions, they make plans, construct cultures, and hold certain values, and their behavior is influenced by such values, plans, and purposes. In short, a human being lives in a world which has "meaning" to him, and, because his behavior has meaning, human actions are intelligible in ways that the behavior of nonhuman objects is not. The opponents of this view, on the other hand, will maintain that human behavior is to be explained in the same manner as is the behavior of other objects of nature. There are laws governing human behavior. An action is explained when it can be subsumed under some such law, and, of course, such laws are confirmed by empirical evidence.

I hold that both of these apparently incompatible views are true. This means, of course, that I shall hold them to be compatible and that I shall hold that the sense of 'understand' in which "understandability" is a unique property of man is not the sense of 'understand' in which to explain something is to understand it.

Let me now turn to Weber's discussion of *verstehen*. Weber begins his explication of the concept by distinguishing two sorts of understanding. He holds that the understanding of an action can be either rational or emotionally empathic. Actions are understood rationally when we "attain a completely clear intellectual grasp of the action elements in the intended context of meaning."[2] They are understood empathically when "we can grasp the emotional context in which the action took place."[3]

These comments are hardly sufficient. Let me begin to unravel them by inspecting some examples of rational understanding. Weber suggests the following examples of the sorts of objects of rational understanding. First, he claims that we can have rational understanding of cases where persons employ mathematically or logically related propositions in reasoning or of cases where persons correctly carry out a process of reasoning according to accepted modes of thinking. Accordingly, we can also rationally understand a person's actions when they are seen to be the result of means-ends reasoning. Here the selection of means (or the reasons for the selection of means) has the same sort of intelligibility as an argument.[4] Weber later suggests that ideas or utterances are also subject to this sort of understanding.[5]

---

[2] Max Weber, *The Theory of Social and Economic Organization* (New York: Free Press, 1947), p. 91.

[3] Ibid.

Now I would suggest that what these examples have in common is that they all refer to types of behavior which are rule governed. Arguments, means-end or otherwise, are governed by logical rules specifying what counts as an acceptable or valid argument. Likewise, utterances are rendered intelligible because they are constructed according to semantic and syntactical rules which specify how the words relate to the world on one hand and how they relate to other words on the other. Without such rules we would not understand what counted as a valid argument, and we would not be able to understand linguistic behavior. Without accepted logical criteria one argument is as good as another, and without semantic and syntactical rules, writing and utterances dissolve into markings and noise.

Thus, to understand an action rationally is to see it as falling under certain rules[6] which specify what counts as adequate performance in a given area and which, thus, lend that kind of activity a sort of intelligibility which the behavior of most objects does not possess.[7] Of course, logical and linguistic rules are not the only sorts of rules there are. Manners, for example, are constituted out of sets of rules which govern certain aspects of our social lives and which lend activities such as eating or attending the theater a kind of intelligibility which they would not otherwise possess. Games, as well, are rule-governed activities, and thus, for example, the existence of the rules of chess is a condition of our understanding a certain situation in the game as being in check or a certain move as castling. It is perhaps unnecessary to add that we understand many cultural objects in the same way. Consider, for example, our understanding of a voting machine or of voting. These objects and acts are not defined in terms of their physical properties, but rather in terms of the set of social conventions which govern them. Thus, to understand voting it is necessary to see the activity as the following of these rules.

---

[4] Ibid.

[5] Ibid., pp. 94, 95.

[6] Although Weber surely would not put the point this way.

[7] It may be helpful here to distinguish between a rule and a scientific law. A rule prescribes how an action *ought* to be done. Rules can be broken without being invalidated. Laws, however, of the sort sought by scientists, describe regularities in phenomena. Thus, if a counterinstance to a law can be found, the law is shown to be false.

As the word 'rational' suggests, Weber has selected the sets of rules presupposed by the construction of arguments as paradigmatic of what is involved in rational understanding.[8] Thus, to rationally understand an action is to grasp the reasons which the agent had for what he did *and* to see them as good reasons. It is important to note here that, given this analysis of rational understanding, the opposite of 'understandable' is not 'inexplicable.' To say that we cannot rationally understand a given action is to say either that it does not fall under any set of rules, e.g., having a muscle spasm, or that it is done in violation of them. It is not to say that we could not explain it. Thus, there is a sense of 'understand' in which we could not understand someone's holding that $2 + 2 = 9$ or acting on some instances of $p > q$, q therefore p, although we likely could explain the occurrence of these events. In Weber's sense of 'understand,' we cannot rationally understand an act when it is irrational, stupid, or purposeless, but this by no means indicates that the act is inexplicable.

Empathic understanding is considered by Weber to be appropriate in those cases where rational understanding cannot be had. Such cases seem to be of two sorts. The first type concerns ultimate values. I suspect that Weber holds that such values cannot be rationally understood simply because they are ultimate; that is, they cannot be held to be worthwhile as means to any more fundamental ends. Therefore, they are not susceptible of rational justification and, thus, cannot be rationally understood.[9] The second kind of action which cannot be rationally understood is irrational action. Weber holds that actions are irrational generally because emotions such as anger, envy, love, and pride cause men to deviate from rational courses of action. Thus, although we cannot have rational understanding of such actions, we can understand them empathically.

---

[8] Weber's explication of an ideal type is developed on this assumption. An ideal type generally specifies what sort of actions would be rational for an agent holding certain values or purposes in certain circumstances. (See Weber, p. 92.) Thus, many of the interpretations of ideal types, e.g., that ideal types can be understood as laws having no instances which can be deduced from true scientific theories, simply are beside the point as far as Weber's position is concerned. See Carl Hempel, "Typological Methods in the Social Sciences," in *Philosophy of Social Science: A Reader*, ed. Maurice Natanson (New York: Random House, 1963).

The objects of empathic understanding seem to be subjective (mental, not nonobjective) states and the connections between subjective states and behavior. Thus Weber writes:

> The more we ourselves are susceptible to them, the more readily can we imaginatively participate in such emotional reactions as anxiety, anger, ambition, envy, jealously, love, enthusiasm, pride, vengefulness, loyalty, devotion, and appetites of all sorts, and thereby understand the irrational conduct which grows out of them.[10]

As the passage also indicates, the method of empathic understanding seems to involve an individual imaginatively putting himself in someone else's place and asking questions like, "How would I feel under such conditions?" or "How would I behave having such feelings?" One can answer such questions, of course, only if he has, in fact, been in the appropriate circumstances or has had the requisite mental states. Thus, one cannot have empathic understanding when he has not had the requisite experiences.

Now, the nature of an adequate explication of the sense of 'understand' as employed in these remarks has been a matter of some controversy. Weber and other advocates of *verstehen* are often supposed to be holding either that one cannot understand the meaning of a word denoting a mental state unless he has had that mental state himself or that one's own experience constitutes adequate evidence for laws connecting mental states and behavior. I shall show shortly, however, that Weber held neither of these views. But, then, in what sense of 'understand' is it the case that we cannot understand an experience we have not had?

Weber holds in the foregoing remarks concerning empathic

---

9. See Weber, *Social and Economic Organization*, pp. 90-93.

10. Ibid., p. 92.

understanding that to be able to understand a mental state or the connection between a mental state and an action is to be able to find a reasonable analogue of the mental state or the mental state connected to the action in one's own experience. What sort of understanding is this?

Let me first consider the understanding of a mental state. The question is, Is there any sense of 'understand' in which we cannot understand a mental state which we have not had? I shall argue (I believe in agreement with Weber) that the answer to this question is, Yes. The most likely illustration of the point here perhaps is the understanding of the sexual feelings of members of the opposite sex. Initially, it has to be noted that not having had such feelings does not seem to present a great difficulty to us. We can talk about such feelings quite sensibly, we can discourse about such feelings quite successfully in explaining behavior,[11] and we generally seem to face no particular difficulties understanding any variety of discourse about such feelings.

What, then, is lacking? Is there any respect in which our understanding of mental states we have not had differs from our understanding of those we have had? I would like to suggest that our position with respect to mental states which we have not had is in some respects analogous to the understanding men possessed of genes from the time at which they were postulated in Mendel's law until the time that they were identified with elements of chromosome material and seen by means of electron microscopes. A gene could be called a *that which*. It is "that which" accounts for the distribution of certain properties within a biological sample. Such a concept is, of course, most functional. It enables us to construct a highly sophisticated theory of the inheritance of biological properties, and it suggests criteria for the direct identification of genes when adequate techniques of observation are available. However, it does not contain anything about the intrinsic properties of genes. Likewise, a mental state which we have not had is conceived of as a *that which*. To put it more precisely, mental states are frequently, if not always, defined in terms of their causal relationships.

---

[11] Accordingly, I am inclined to think that the sort of understanding involved in emphatic understanding is of little interest except as a heuristic device for generating hypotheses. See discussion that follows on pages 34-35.

Such a state is "that which" causes a certain sort of behavior in appropriate circumstances.[12] Such a conception is sufficient for our discourse about mental states. But there is a kind of knowledge of mental states which can only be gotten by being in them. We only learn what a mental state "looks like" by having it. For example, we may be perfectly well able to talk about a pain of a certain sort while never having had it. However, we cannot know what a pain "feels like" unless we have had it. Since we do not and cannot have direct access to the mental life of another person, we can only learn the "feel" of a mental state from our own case. Thus, there is a kind of knowledge of mental states which can be had only by being in them. If the intrinsic properties of mental states can be identified with their "feel," then one can understand only the intrinsic properties of mental states which he has had. I would suggest that Weber's position that, while we may significantly talk about mental states we have not had, we can empathically understand a mental state only if we have had a similar one, can best be regarded as an implicit recognition of this point.

In what sense is having experienced some connection between a mental state and an action requisite for empathically understanding that connection? Here I would suggest that the kind of understanding involved is simply familiarity. Suppose, for example, we knew someone who smiled whenever he was angry. Suppose we knew that he was not merely forcing a smile to cover his anger, nor was he being sarcastic. Rather he simply and regularly responded to anger with a smile. Grant for the sake of the argument that this is both logically and empirically possible. I suspect we might maintain that we simply did not understand such a person. Here we would mean not that his behavior was inexplicable, or irrational, but rather only that it was foreign to our experience. It is not what we could do. Therefore, it is not something we can readily understand. Thus, to empathically understand the connection between a mental state and an action is to see the connection as part of one's own experience. It is to feel, "That is what I would do if I felt that way."

---

[12] My view of the logic of mental states is similar to that of Armstrong. See D. M. Armstrong, *A Materialistic Theory of the Mind* (New York: Humanities Press, 1968), pp. 82-85.

Two points must be made concerning the distinction between empathic and rational understanding. First, given the above expansion of Weber's views, claims of rationally understanding an action can be assessed according to public and objective criteria. However, this is not the case with respect to empathic understanding.

I have attempted to explain the notion of rational understanding by pointing out that such judgments about actions depend on the action's being seen as the following of various sorts of rules. But such rules are the conventions accepted by a community which prescribe acceptable or appropriate behavior in a certain domain.[13] Rules and rule-following acts are thus public.

On the other hand, in Weber's view a claim to empathically understand someone's behavior is not so much a claim about the character of the behavior as it is a comment about the previous experience of the maker of the judgment. If Jones correctly claims to have empathic understanding of Smith's emotional states, what we learn from this is (a) that Smith has certain emotional states (this judgment does not require empathic understanding) and (b) that Jones has had the previous experience necessary to enable him to have empathic understanding of Smith. Thus what Jones's claim of understanding Smith adds to the fact that Smith is in a certain mental state is a fact about the previous experience of Jones and not an additional fact about Smith. Thus, although the kind of phenomenon involved in emotional understanding may add considerably to the quality of our interpersonal relations, it does not appear to be necessary to the success of the social or behavioral sciences except to the extent that our own experience can serve a heuristic function in generating hypotheses.

Second, although most of the discussion and criticism of Weber's views have centered around the notion of empathic understanding, rational understanding is clearly the more central concept in Weber's methodological thought. An ideal type, for example, seems to be nothing more than a description of a rational pattern of behavior given the values, purposes, circumstances, and conventions of a community. Empathic understanding is put forth by Weber as a device whereby aberrations from the indicated course of action can be grasped. As I have indicated, I find Weber's remarks on empathy to be scientifically uninteresting and—as they seem to have been the source of much confusion in the literature and to have obscured what is valuable in Weber—unfortunate. I shall, therefore, direct the larger portion of the remainder of this paper to issues concerning rational understanding.

---

[13] As Weber points out, the issue for the sociologist is not the correctness of society's conventions. In Weber's terms, to call an act rational is to say nothing more than that the act conforms to the accepted set of rules governing the activity. Such a judgment does not express approval of the rules themselves. See Weber, *Social and Economic Organization*, pp. 89-90.

In order to complete the discussion of the notion of *verstehen*, it is necessary to consider the relations between the concepts of understanding and explanation. Consider the following passage:

> A motive is a complex of subjective meaning which seems to the actor himself or to the observer an adequate ground for the conduct in question. We apply the term 'adequacy on the level of meaning' to the subjective interpretation of a coherent course of conduct when and insofar as, according to our habitual modes of thought and feeling, its component parts taken in their mutual relation are recognized to constitute a 'typical' complex of meaning. It is more common to say 'correct.' The interpretation of a sequence of events will on the other hand be called *causally* adequate insofar as, according to established generalizations from experience, there is a probability that it will actually occur in the same way. . . .
>
> A correct causal interpretation of a concrete course of action is arrived at when the overt action and the motives have been correctly apprehended and at the same time their relation has become meaningfully comprehensible.[14]

The crucial feature of this passage is the distinction made between a correct interpretation of meaning and a causally adequate interpretation. Weber makes clear that there are two separate judgments to be made concerning an action, one concerning its meaningfulness and the second concerning its explanation. Perhaps the point can be restated in the following way. People normally behave for reasons; that is, their behavior is frequently explained in terms of what they want and what they believe about how to accomplish their purposes. Let me call the set of reasons that a person holds for what he does the "rationale" for his action. I shall use the letter R to stand for a rationale. Let us suppose that we explain some action, A, by R. Now we can ask two quite different questions about the relation between R and A. We can ask:

---

[14] Ibid., pp. 98, 99.

(1) Is R the explanation of A? or
(2) Does R justify A? [15]

The first question asks for the cause of A. Is it in fact the case that A was done because of R? Weber is quite clear that verification of a claim that R causes A requires the empirical confirmation of a generalization connecting R and A. The evidence for a claim that R justifies A (or, in Weber's terms, for a claim that R is adequate on the level of meaning) is quite a different matter. Here it is not a matter of the empirical confirmation of some lawlike proposition, but rather of determining whether the action or its rationale meets accepted criteria of evaluation. Was that action done correctly? Were the reasons good ones? Deciding whether a rationale justifies an action is more like determining the validity of an argument than empirically verifying a generalization. One assesses the rationale according to accepted standards of judgment. Thus, generalizations which are not confirmed by the evidence are judged to be false; however, rationales which do not conform to acceptable criteria of judgment are unintelligible.

It is also the case that judgments concerning responses to question (1) do not entail anything about judgments concerning responses to question (2) and conversely. It does not follow from the fact that R explains A that R justifies A, and it does not follow from the fact that R would justify A that R explains A. (Of course, if we claim that we understand A because it is explained by R which is rationally intelligible, it must be the case that R does explain A. It does not follow from this, however, that "R justifies

---

15. I have elected to use the phrase "Does R justify A?" instead of the more cumbersome "Does R render A meaningful?" in order to simplify the discussion. Let me remind the reader that Weber's paradigm of rational understanding is the sort of intelligibility exhibited by arguments. To say that an action is rationally understandable is frequently to say that given the values and circumstances of the agent his action is justified. The notion of rational understanding is broader than the notion of justification, including the sort of intelligibility exhibited by acts which follow various social rules, e.g., voting. The task of sorting out the various types of rational intelligibility which might be involved in Weber's notion is an important one and seems to have been rather neglected. However, so far as I can see, nothing in my subsequent argument turns on making these distinctions. Thus, the shift from 'meaningful' to 'justified' seems appropriate.

A" is an empirical claim.) Thus, Weber holds that empirical evidence is required to verify an explanation even if there is adequacy of meaning.

> . . . even the most perfect adequacy on the level of meaning has causal significance . . . only insofar as there is some kind of proof for the existence of a probability that action in fact normally takes the course which has been held to be meaningful.[16]

Likewise, if we find that R does not justify A, what follows is not that A is inexplicable but that it is rationally incomprehensible.

> If adequacy in respect to meaning is lacking, then no matter how high the degree of uniformity and how precisely its probability can be numerically determined, it is still an incomprehensible statistical probability whether dealing with overt or subjective processes.[17]

Now, it is my contention that it is both the force and merit of Weber's doctrine of *verstehen* that it recognizes the distinction between questions of type (1) and (2). That is, Weber holds that men do behave for reasons, that reasons both explain and justify actions, and that whether a reason explains and whether it justifies are different questions. To subsume an action under some empirical law is to explain it, while to see it as warranted by acceptable reasons is to understand it. I shall try to show in the next portion of this paper that most of the critics of the doctrine of *verstehen* miss this point.

The critics of the above notion of *verstehen* can be divided into three groups which I shall call radical empiricist, moderate empiricist, and antiempiricist.

The radical empiricist is inclined to dismiss any explanatory scheme employing mental states on the grounds that at best such states must be inferred from behavior and at worst they are entirely incapable of empirical identification.[18] The restrictions placed on the introduction of explanatory concepts by the radical empiricist are sufficiently demanding that mental states cannot meet the criteria. It is often demanded that any concept employed in the explanation of behavior refer to something directly observable or be analyzable into objects or events directly observable. This criterion is held to be too restrictive for mental concepts to meet. Thus, the radical empiricist is led to disavow the use of terms denoting mental states in his explanatory framework and to embark on the construction of a behavior theory without mental states. Since the notion of *verstehen* depends on the view that mental states are involved in the genesis of behavior, the radical empiricist will dismiss the position out of hand.

---

[16] Weber, *Social and Economic Organization*, pp. 99-100.

[17] Ibid., p. 100. It is worth noting that the issue of whether an action can be understood is not the question of whether mental states are involved in its explanation. This is a necessary condition of rational understanding, but it is also required that these mental states embody an acceptable argument.

[18] Skinner at times seems to hold such a position.

I am inclined to believe that such a position is simply absurd. It depends both on an inadequate view of concept formation and an overly restrictive concept of experience. At this point let me note only that an adequate view of experience must recognize our acquaintance with our own subjective world as a type of experience and that once this is recognized it follows that the radical empiricist's view of mental states is empirically false.

The moderate empiricist is generally willing to employ mental concepts in explaining behavior. Thus, he is willing to talk about motives, purposes, values, and the like, insisting only that such concepts be tied sufficiently to experience to render them susceptible to intersubjective identification and that explanations employing such concepts be capable of empirical verification. The conflict between the moderate empiricist and the *verstehen* advocate consists in the fact that the *verstehen* advocate is frequently held to deny both of the above points.

Specifically, Weber and other advocates of *verstehen* are thought to hold (a) that we can understand the meaning of terms denoting mental states only if we have had such states ourselves and (b) that the empathic or imaginative identification of one person with the experience of another constitutes sufficient evidence for a claim about an explanation of some action.[19]

I shall not quarrel with the arguments employed to object to such views. They seem to me to be unexceptionable.[20] Rather, let it suffice to point out that Weber held neither of the above positions and that neither view is required to maintain the notion of *verstehen* as I have characterized it. Besides the fact that Weber specifically disavowed the first position, he frequently talks in such a way as to make clear that failure to empathically understand some mental state does not render us unable to refer to such a mental state significantly or accurately. It has already been made clear that Weber does not hold the second view and that discussion need not be repeated.

---

[19] It is the case, however, that some advocates of *verstehen* do hold such views.

[20] For a discussion of such issues see Ernest Nagel, *The Structure of Science* (New York: Harcourt, Brace & World, 1961), pp. 473-85; idem, "Problems of Concept and Theory Formation in the Social Sciences," in Natanson, *Philosophy of Social Science*; and Theodore Abel, "The Operation Called *Verstehen*," *American Journal of Sociology* 54 (1948): 211-18.

Of more interest is the question as to why such views should be ascribed to Weber when he seems clearly not to hold them. If I may hazard a guess, it would go something like this. The difficulty stems from the fact that many of Weber's critics fail to note the distinction he makes between 'explaining' and 'understanding' an action. Thus, when Weber holds that actions are understood when they are seen to conform to rational criteria or when he holds that by imaginatively projecting oneself into another person's experience we can understand his emotions and their connections with his behavior, he is taken to be advocating an antiempiricist view about the logic of concept formation or about the logic of confirmation. Weber does, of course, say such things about understanding actions. However, he does *not* say them about the explanation of actions. His critics fail to note this because they take his remarks about understanding to be remarks about explanation.

Most interesting among Weber's critics are the advocates of a position which holds that Weber's account of *verstehen* is for the most part adequate, but which maintains that Weber was wrong in demanding empirical evidence for explanation in sociology and in advocating a concept of explanation modeled after explanation in the natural sciences. Sociology, they claim, cannot be understood on a natural-science model. Its patterns of explanation are quite different from a natural-science model and accordingly its canons of evidence are also quite different. The foremost exponent of this position has been Peter Winch. Since his book *The Ideal of a Social Science* provides a detailed critique of Weber's thought, I shall deal exclusively with it.

Winch's account of the concept of understanding is developed along lines similar to those I employed in elucidating Weber's position.[21] Winch's strategy is to develop an analogy between understanding the meaning of an utterance and understanding the meaning of an action. How is it, Winch asks, that English-speaking people are able to use words like 'Mount Everest' to talk about a particular object in the Himalayas? Basically, Winch's response is that a word acquires a meaning because it is involved in a context of rules which prescribe the connection between the word and the world and, thus, prescribe the correct use of the word, i.e., the circumstances, occasions, or objects to which it can be correctly ascribed. In short, to understand the meaning of a word is to know the rules for its correct employment. The fact that 'Mount Everest' is involved with such a set of rules is what turns the physical marks on a piece of paper into a word denoting an object in the Himalayas.[22]

Similarly, Winch argues, it is the fact that some piece of behavior is done according to various social rules which gives it meaning and which renders it an object of understanding. What, for example, is the difference between hitting a ball over the fence with a stick and hitting a home run, or what is the difference between pulling levers in a booth and voting Republican? The difference in each case is that in the second instance, but not in the first, the behavior is seen as the following of some set of rules. Thus, it is the rules of baseball which give hitting a ball over a fence with a stick the import of hitting a home run, and it is the conventions surrounding elections which enable pulling a lever to mean voting Republican. The rules governing such events enable appropriate physical movements to have a meaning, just as linguistic rules enable words to have a meaning. And the fact that rule-governed actions have meaning in this sense indicates that they can be understood in a way quite different from the sort of understanding that the natural sciences provide us of the behavior of physical objects.

Winch infers from the above position that explanation in the social sciences cannot be anything like explanation in the natural sciences. Explanation in the natural sciences is held to be basically a matter of establishing causal sequences between two or more variables. When we hold that C causes E, we hold that a law to the effect that E always follows C in similar circumstances is true. We are held to have explained E when we can deduce it from a conjunction of C-causes-E and C, and we are enabled to predict E when we know C-causes-E and when we know C.

---

[21] To give credit where credit is due, Winch's account has strongly influenced my reading of Weber.

[22] Peter Winch, *The Ideal of a Social Science* (New York: Humanities Press), pp. 24-32.

Winch suggests two arguments to show that explanation in the social sciences is nothing like this. First, Winch argues:

> Now if the position of the social investigator . . . can be regarded as comparable . . . with that of the natural scientist, the following must be the case. The concepts and criteria according to which the sociologist judges that, in two situations, the same thing has happened, or the same action performed, must be understood in relation to the rules governing the sociological investigation. But here we run against a difficulty; for whereas in the case of the natural scientist we have to deal with only one set of rules . . . here what the sociologist is studying . . . is a human activity and is therefore carried on according to rules. And it is these . . . which specify what is to count as 'doing the same kind of thing' in relation to that kind of activity.[23]

I take it that the point of this somewhat opaque passage is this: When the social scientist claims to understand an action, he is grasping its meaning by seeing it as governed by a particular set of rules. But seeing an act as the following of a rule is very different from seeing it as an instance of some regularity or law. Rules can be broken, laws cannot. Rules give meaning to action, laws merely relate the condition under which they occur. Thus, to understand an action is not at all like giving a scientific *explanation* of it.

Winch's second argument concerns the context in which reasons are given for some action. Suppose that a university lecturer cancels his lecture giving as his reason his intent to travel to London. How are we to understand such an explanation, i.e., how does the explanation explain? According to Winch, the lecturer

> . . . does not offer his reason as evidence for the soundness of his prediction about his future behavior. Rather, he is justifying his intention. His statement is not of the form: "Such and such causal factors are present, therefore this will result" nor yet of the form: "I have such and such a disposition, which will result in my doing this"; it is of the

---

[23] Ibid., pp. 86-87.

form: "In view of such and such considerations this will be a reasonable thing to do."[24]

The argument might be put as follows: Reasons provide the justification of an action. Causal explanations do not justify actions. Therefore, reason explanations are not causal explanations. Again the point is that when we make an action intelligible by providing reasons for it, we are not doing the same thing as when we explain an action by deducing it from a law and a set of initial conditions.

Given this position concerning the understanding and explanation of behavior, Winch is led to take a highly critical view of Weber's insistence that sociological explanations are causal and are, therefore, in need of empirical confirmation. Winch states the following argument against Weber's supposed tendency to talk as though the ability to formulate statistical laws and predict behavior were the ultimate test of a claim to understand an action.

> . . . we might well be able to make predictions of great accuracy . . . and still not be able to claim any real understanding of what those people were doing. The difference is precisely analogous to that between being able to formulate statistical laws about the likely occurrences of words in a language and being able to understand what was being said by someone who spoke the language. . . . 'Understanding,' in situations like this, is grasping the point or meaning of what is being done or said. This is a notion far removed from the world of statistics and causal laws . . .[25]

Two arguments seem implicit in these remarks. The first argument maintains that since it is possible to subsume an action under a statistical regularity and generate accurate predictions concerning it without understanding it, it follows that understanding could not depend on the existence of such a regularity. This argument is an obvious nonsequitur, for even if Weber held that understanding an act depended on discovering some such regularity, it is clear that noting this regularity would be only a necessary but not a sufficient condition of understanding the action. However, this being the case, it would simply not follow that because one could subsume an act under some regularity without understanding it that understanding it did not depend on the existence of such a regularity.

[24] Ibid., p. 81.

[25] Ibid., p. 115.

The basic thrust of Winch's argument, however, consists in again noting the marked logical difference between an action having a meaning because it is rule-governed and explaining an action by relating it to some empirical law. As Winch notes, it is the same difference as that between understanding a sentence and being able to predict the occurrence of certain words. The difficulty with these arguments is not Winch's analysis of understanding, but his failure to mark the distinction between understanding an action and explaining it. Thus, although Winch is quite right in arguing that understanding an action is not at all like providing a scientific or causal explanation of it, he is not entitled to infer from this that scientific or causal explanations of action are logically inappropriate in explaining action. Understanding an act is not like a scientific explanation because it is not like an explanation at all. But this is quite compatible with Weber's position that acts are explained by showing them to fall under laws.

Let me try to make my objection to Winch clear by looking at some aspects of his arguments. At one point Winch maintains that because what a man is doing when he gives his reasons for his actions is justifying his actions, it follows that he cannot be providing a causal explanation of his actions. Quite true, but when a man is justifying his actions, he is not explaining them at all. It does not follow from the fact that some act is justified that it will occur. What accounts for the occurrence of an action is the fact that someone accepts the justification. However, believing that an act is justified and its being justified are two different things. Beliefs are mental states. They are suitable candidates for causes of behavior. That an act is justified is not a comment about someone's mental state. It is a comment about the logical properties of an argument, i.e., that it is a good argument. Thus, Winch's problem is not his view that the justification of an act is not a causal explanation of it but his view that the justification of an act explains it.[26]

---

[26] Behind this problem lies an ambiguity in the word 'reason.' Suppose Jones acts because he believes that *p*. Now it is *p* which if true would justify Jones's action. However, it is believing *p* which would explain it. Both *p* and believing *p* may be called the reason for Jones's action. I suspect that ultimately the claim that the justification of an act explains it stems from a failure to note this ambiguity.

Suppose we consider Winch's example in his argument against Weber. Here Winch claims that understanding the meaning of an action is like understanding the meaning of a sentence. Again, I do not wish to dispute this. However, it must be pointed out that understanding the meaning of a sentence and explaining it are two different matters. It may be the case that one does not understand a sentence when he has discovered statistical laws about the occurrence of sentences or words, but he has not explained it when he has understood it either. Explaining the occurrence of a sentence may require understanding it. It is after all unlikely that people will often utter strings of meaningless noises, but explaining an utterance is first and foremost a matter of accounting for its occurrence, knowing why this utterance and not another was made. In short, explaining an utterance is a matter of discovering the conditions under which an utterance will occur, i.e., knowing its causes. Again, Winch's account of understanding is incompatible with the position that human actions are susceptible of causal explanation only because Winch fails to adequately distinguish between understanding an action and explaining it.

My response to Winch's objections to Weber ought by now to be clear. Winch takes remarks by Weber about the nature and confirmation of *explanations of behavior* to be remarks about the nature of *understanding an action* because he misses the distinction Weber makes between understanding an action and explaining it. His criticisms of Weber, therefore, are based upon an exegetical mistake. More important, however, is the fact that the exegetical mistake is precipitated by a philosophical one, the failure to distinguish between understanding and explaining an action.

Allow me at this point to back off and take a broad and programmatic look at this discussion. What I hope to have accomplished here, in addition to having pointed out a rather important distinction between explaining and understanding, is to have provided some reason to believe that a commitment to treating man as an object of scientific scrutiny does not require us to give up the more traditional view of human nature. What I find attractive in Weber's account of the explanation and understanding of man is that is succeeds in reconciling these apparently divergent strains. Man, for Weber, is a purposeful, rational, and value-seeking thing.

Man has a subjective life. However, man's behavior can be accounted for scientifically by the simple expedient of incorporating man's subjective life into an account of his behavior. What one has to realize is that when we note that human actions have meaning and that we often explain actions by noting the reasons an agent had for doing them, we open up other ways of understanding men which, while they are not scientific, are perfectly compatible with a scientific outlook. Thus, we can treat man as a rational animal and as an object of science without contradiction.

Perhaps the impact of this reconciliation can be made clear by one more reference to Winch's argument. Suppose it were the case that an adequate doctrine of understanding was incompatible with a scientific account of behavior. Two conclusions are possible. First, we may conclude, as does Winch, that human action is not an appropriate object of scientific explanation. However, one might also conclude that, since scientific modes of explanation are sacrosanct, we will have to revise our opinions of man accordingly. If, in fact, the notion of *verstehen* is tied to a view of man as a rational, purposeful, rule-following agent, and if this is not compatible with treating man as an object of scientific scrutiny, then we will simply have to stop thinking of man as a rational, purposeful, rule-following agent.

I am inclined to think that for very many in and out of the academic community it is this second conclusion which will be drawn rather than Winch's. But what is at stake here is nothing less than the maintenance or the erosion of the distinction between persons and things. To decide that man is not a rational or purposeful agent is to decide that he is much more like other physical objects than he has previously been supposed to be. The difference between man and rocks and sticks is merely one of complexity. Man surely does not deserve a category of his own.

Unfortunately, our categories have a way of catching up with us. We cannot begin to erode the distinction between persons and things without beginning to feel that men ought to be treated like things. We ought not to forget that very many of our ethical, political, and educational views are inextricably tied to a traditional view of human nature. Men are supposed to be treated as ends rather than means because they are persons, not things. Many of the philosophical arguments which undergird our civil liberties presuppose that man is a rational agent. It goes without saying that decisions in education concerning the development of children's intellectual capacities, respecting their freedom of choice, and initiating them into democratic values will depend in large measure on the view we hold of the nature of man. If we decide that men really are not significantly different from other things, we may be led to conclude that they are proper objects of manipulation and control and that arguments which end with conclusions about freedom, reason, democracy, and the like are nothing but the dreams of fuzzy-headed romantics or philosophers.

I am most alarmed to see developing in American thought a battle between "proscience" and "antiscience" culture, the former talking with enthusiasm about human engineering, programmed instruction and programmed people, and reinforcing socially desirable behaviors, while the latter talks about freedom, authenticity, and the oppressive nature of scientific thought. One has merely to juxtapose a reading of two books like Skinner's *Walden Two* and Roszak's *The Making of a Counter Culture* to see the tension in living color. If my argument is correct, these are not the choices. It will be tragic if American culture and American education fight this battle when the sides are based on bad philosophy.

Note: References not cited in the footnotes of this chapter include Donald Davidson, "Actions, Reasons, and Causes," *Journal of Philosophy* 60 (1963): 685-700, reprinted in *Readings in the Philosophy of Science*, ed. May Brodbeck (London: Macmillan & Co., 1968); Edward Madden, ed., *The Structure of Scientific Thought* (Boston: Houghton Mifflin Co., 1960); Jane R. Martin, *Explaining Understanding, and Teaching* (New York: McGraw-Hill Book Co., 1970); Alfred Schultz, "Concept and Theory Formation in the Social Sciences," *Journal of Philosophy* 51 (1963): 257-73; idem, *The Phenomenology of the Social World* (Evanston, Ill.: Northwestern University Press, 1967); Max Weber, *The Methodology of the Social Sciences* (New York: Free Press, 1949).

CHAPTER III

Theories Are Tested by Observing the Facts: Or Are They?

HUGH G. PETRIE

I

"The first rule that we must be prepared to accept as we judge the relative merits of different [learning] theories is this: *All the theorists accept all of the facts.*" [1] This view, enunciated by one of the most respected learning theorists, is, I believe, a view which is held almost without question by most psychologists, educators, and laymen. It is also, I believe, a view which is a most profoundly mistaken one. It is the purpose of this essay to indicate briefly why this view is mistaken and to show with a discussion of several examples what a pernicious influence this view has had on the development of psychological learning theory and most especially on the application of psychological learning theory to educational concerns.

The supposition that all the theorists accept all the facts is not a very surprising one. It is a fairly common piece of current scientific folklore and just a simple restatement of the generally accepted positivist principle that we can always draw a sharp and clear distinction between an observation language which reports the facts of our environment and a theoretical language which interprets those facts. Thus the presupposition is that there is a neutral data language upon which all agree and differing theoretical languages interpreting the data over which there can be disagreement.

---

[1] Ernest R. Hilgard and Gordon H. Bower, *Theories of Learning*, 3d ed. (New York: Appleton-Century-Crofts, 1966), p. 9.

The initial investigation underlying this essay was supported in part by a small project grant from the Office of Education, Contract 0-8-080023-3669(010).

And yet, there has recently arisen a serious challenge offered by such men as Hanson,[2] Quine,[3] Toulmin,[4] Kuhn,[5] and Petrie[6] to such "obvious" presuppositions. These men have begun to argue that scientific theories are radically underdetermined by experience and that although scientific theories must have empirical content—be testable by experience—they do not and cannot arise solely out of experience. It has been argued that what even counts as experience is essentially theory dependent. That is, two scientists may look at the "same" thing and, because of different theoretical perspectives, may literally not *see* the same object. What is relevant for one theory may be logically and methodologically incapable of being observed under the presuppositions of a different theory.

It should be emphasized at this point what a truly radical conception this is. It might be supposed what is being claimed here is that any science in fact focuses on certain features of experience to describe and ignores others. For example, classical physics, it has often been said, owes much of its success to having concerned itself with just the right physical properties, position and momentum, ignoring color, taste, and so on. If this is the sort of thing being claimed, then why all the fuss?

But this would be to miss the point entirely. For such a "focusing" conception of science logically presupposes a kind of neutral experiential base upon which one may focus, now here, now there. Correlatively, a neutral observation language is also presupposed within which one could in principle describe all the physical properties of situations and events, leaving it to the scientific theory to pick out those features which are to be covered by the theory. The nonfavored features are still "there"; they are simply not deemed relevant.

---

[2] N. R. Hanson, *Patterns of Discovery* (London: Cambridge University Press, 1958).

[3] W. V. O. Quine. *World and Object* (New York: John Wiley & Sons, 1960).

[4] Stephan Toulmin, *Foresight and Understanding* (New York: Harper Torch Books, 1961).

[5] T. S. Kuhn, *The Structure of Scientific Revolutions* (Chicago: University of Chicago Press, 1962).

[6] Hugh G. Petrie, "The Logical Effects of Theory on Observational Categories," Office of Education Contract 0-8-080023-3669(010).

However, it is the position of the view under consideration that no such neutral observation language exists nor can experience be described independently of theory. Thus, under this radical view the very categories of things which comprise the "facts" are theory dependent—the exact opposite of Hilgard's optimistic claim.

It should be pointed out that the discussion thus far has proceeded and will continue to proceed on the level of the *description* of experience. That is, I have taken what has been called the linguistic turn.[7] Now it has been argued that such a move is ultimately unacceptable.[8] According to this line of criticism I should not even speak of our *experience* of the world, already a transcendental turn, but rather of the world itself if, indeed, I wish to say something about the world. I cannot here enter into this very interesting discussion. My reason for avoiding it is that when Hilgard says that all the theorists accept all the facts, he clearly means that psychologists do not argue over the experimental data which is reported in the journals. Clearly this is on the level of the description of the psychologists' experience. Thus I will also not here worry over the exact ontological status of the troublesome concept of a fact.[9] Rather, I shall content myself with treating as a fact those linguistic entities which appear as "data" categories in the typical experimental psychology literature. These are clearly the facts to which Hilgard is referring and they are at least closely related to the "facts" that philosophers discuss.

Thus when I speak of "observational categories" I have in mind such things as "trial," "bar press," and "series of nonsense-syllable responses." These are the kinds of terms which appear in journal articles and which count as "observable" in some sense or other. On the other hand, when I speak of "what we see," I mean the experiences by virtue of which we go on to apply the foregoing observational categories.

---

[7] R. Rorty, ed., *The Linguistic Turn* (Chicago: University of Chicago, 1967).

[8] Henry Veatch, *Two Logics* (Evanston, Ill.: Northwestern University Press, 1968).

[9] See, for example, George Pitcher, ed., *Truth* (New York: Prentice-Hall, 1964).

II

An extreme form of the theory-dependency thesis of observation would present us with a most radical kind of Whorfianism. For if each of us sees only what the theory we have enables us to see, and it is furthermore granted that everyone's conceptual scheme differs at least slightly from everyone else's and, finally, that our conceptual schemes are, in some sense, our theories of the world, then no one ever sees precisely what anyone else sees, and a rigorous notion of intersubjective confirmation or justification of some one theory as over against another is logically out of the question. Such an extreme view often seems to be implied by some of the things Kuhn says.[10]

I do not think that such an extreme view is correct. For one thing, it has all the difficulties which any radical skepticism faces, in addition to some of its own. First of all, if this kind of theory-dependency thesis is even intelligible at all, it will be intelligible *on its own grounds* only in terms of some theory which determines observational categories sufficient for us to see the intelligibility of the theory-dependency thesis. It seems obvious that such an all-embracing metatheory is nothing more nor less than philosophy and thus that philosophical argumentation is appropriate to the theory-dependency thesis. For if the thesis actually asserts that it itself is outside the realm of any justification, even a philosophical justification, then quite clearly we can have no justification for accepting it, and yet equally clearly the thesis is capable of being argued about.

Second, even if we were to grant the extreme Whorfian version of the theory as a metaphysical possibility, we could not on epistemological grounds ever assert or deny this possibility. For as Quine has so adequately pointed out,[11] there is no way of deciding on the basis of the empirical evidence between someone's looking at the world in a radically different way and a mistake in translation. To make sense of the differences in conceptualization we do find, we must assume a tremendously large core of common conceptualization as a background.

But if such an extreme construal of the theory-dependency thesis is untenable, nevertheless there remains a version of the thesis which is of vital importance to the relationship between epistemology and learning theories. For the traditional ground of objectivity for all empirical theories and learning theories in particular—viz., appeal to observation—is itself asserted to be theory-bound. But this means that the role of observation in providing all or even part of the objective basis for theory stands in need of radical reanalysis. One cannot *simply* appeal to observation to settle empirical disputes. A theory might perhaps be refuted "objectively" on rationalistic grounds. The point is that such objectivity would be inconceivable to those who, like Hilgard, look upon observation as *the* objective ground of theory.

---

[10] *Structure of Scientific Revolutions*, chap. 10.

[11] *Word and Object*, pp. 73-79.

On the other hand, the rejection of the extreme subjectivist interpretation of the theory-dependency thesis places me somewhat in line with at least the thrust of Hilgard's remarks. For if it is not true that any theory is as good as any other, then there must be *some* way or other of providing an objective ground for choosing between them even if this ground is not a neutral observation language. What is right, therefore, about Hilgard's remarks and what should be recoverable in an appropriate reanalysis is that theories are *commensurable* in some way or other.[12] In a way, it is the purpose of the rest of this essay to begin exhibiting in use varying modes in which we can objectively compare theories without appeal to a neutral observation language, although, of course, on investigation *some* sort of appeal to observation reanalysed will be made. For certain purposes and within the bounds of certain presuppositions, it is true that observation grounds empirical knowledge. For this reason I will occasionally lapse into language reminiscent of Hilgard. I can only hope that such lapses will be clearly justified by my having made explicit my purposes and presuppositions.

I should also note here that even a mild form of the thesis of the theory dependency of observation has a paradoxical sound to many ears. For example, it might well be urged against some of the things I will say in the sequel that these remarks are themselves theory dependent and hence not compelling for anyone who does not accept my theories. In reply, let me repeat that I think such an objection has weight only against the extreme form of the theory-dependency thesis. And I have already rejected that form of the thesis. In addition I have, like Quine, rejected the possibility of a philosophical justification of epistemology.[13] And, like Quine, I will make free use of science in treating problems of "epistemology," or "critical methodology," or whatever you will. Rational discussion can still take place against the background of the presumptions of philosophy and science. It is only if it is thought that my remarks are an attempt to justify this background a priori in a way which is independent of the theory-dependency thesis that misunderstanding might arise. I make no such sweeping claims. The scientific and other results that I use *are* subject to the theory-dependency thesis. But, since I have not ruled out the possibility of some of these results being "better" than others, my position results in no contradiction.

---

[12] See, for example, I. Lakatos and A. Musgrave, eds., *Criticism and the Growth of Knowledge* (Cambridge: Cambridge University Press, 1970) for a recent discussion of the problem of commensurability as raised by Kuhn's work.

But now what reasons are there for believing even the modified version of the theory-dependency thesis to the effect that objectivity is not solely guaranteed by neutral observation. The reasons are complex and varied and can be found discussed at great length in some of the authors already cited, e.g., Kuhn, Hanson, Quine, Petrie, and Toulmin.[14] For my purpose here I wish briefly to run over two lines of argument. On the one hand, there is the apparent failure of all programs to identify the basic particulate materials of observation. In philosophy, this failure is exemplified by the downfall of sense-data theories and phenomenalism as viable answers to the problems of perception.[15] In psychology, the work of gestalt psychologists in perception bears ample testimony to the importance of individual conditions on what is perceived.[16] Various experiments on "set" also indicate the extreme difficulty if not impossibility of finding a unique "neutral" or "objective" description of observation. The very categories of what can be perceived seem theory dependent.[17]

---

[13] W. V. O. Quine, *Ontological Relativity* (New York: Columbia University, 1969), especially the chapter entitled "Epistemology Naturalized."

[14] One might also consult Israel Scheffler, *Science and Subjectivity* (Indianapolis: Bobbs-Merrill Co., 1967) for an "opposing" viewpoint which, nevertheless, seems to grant the major points of the theory-dependency thesis.

[15] See, for example, G. J. Warnock, ed., *The Philosophy of Perception* (Oxford University Press, 1967).

On the other hand, if there are no basic perceptual categories to which we can refer observation or into which observation could be analyzed, then it would seem that the criteria for membership in a purportedly "basic" perceptual category must be more or less indefinite and amenable to modification. One could avoid this looseness only by giving up on classification and reverting to simply proper naming. But once that stage is reached, the very idea has been abandoned of finding some public, repeatable, intersubjectively verifiable observation that could serve as the objective basis for theoretical interpretation. As in sense-data theories, complete specificity is purchaseable only at the price of being unable to say anything *about* the specified item.

III

These two lines of argument come together in a most illuminating way in a consideration of operational definitions. Operational definitions are crucially relevant to a discussion of the theory-dependency thesis of observation because it is just here that most theorists of Hilgard's persuasion take their stand on the distinction between the facts and what can be inferred on the basis of the facts. Operational definitions are the point at which theory makes contact with objective experience, and yet, as I shall urge, taking the theory-dependency-of-observation thesis seriously leads to a relativization of operational definitions which is in conflict with their purported role as objective anchors of theory in experience. Furthermore, this necessary relativization of operational definitions helps explain a number of current methodological problems in learning theories.[18]

---

[16] See, for example, K. Koffka, *Principles of Gestalt Psychology* (New York: Harcourt, Brace & World, 1935) and W. Kohler, *Gestalt Psychology* (New York, Liveright, 1947).

[17] See also M. Herskovits, D. T. Campbell, and M. Segall, revised by Segall and Campbell, *A Cross-Cultural Study of Perception* (Indianapolis: Bobbs-Merrill Co., 1969).

Although there are many conceptions of what an operational definition should be and at least as many problems relating to it as there are conceptions, I shall here need only a very general notion of an operational definition. I shall consider an operational definition to be the association of a definite, observable, testing operation with the term being defined as its (the definiendum's) criterion of application. This formulation should suffice for my purposes.

The first thing to notice is that in order to establish the reliability of the operational definition, one must be able to identify both occasions of the application of the testing operation and what happens in each instance as a result. For example, one must recognize any number of occasions of the taking of an intelligence test if one is using such an operational definition of intelligence. Although these events all fall under the concept of taking the test, they also differ among themselves in an indefinite number of other characteristics or else they would not be separable events. This is the exemplification of the requirement, just discussed, that the events be classified into some category or other and not just named. What this also shows is that there are two distinct logical roles which must be played in an operational definition. These are: first, the role of observable terms, namely, the test operation and its result, and, second, the role of theoretical terms defined by the operational definition. The observable terms must be assumed to be reliably identifiable from occasion to occasion. For it is only if these terms are identifiable that we can go on to ask the further empirical question of whether or not these two observables go together. We must assume that there are some persisting, underlying properties of one form or another by virtue of which we can classify two or more events as being events of the kind determined by the properties. Thus the problem of operational definitions is not that of the distinction between observables and underlying properties. *All* operational definitions assume some sort of underlying properties as the basis of being able to apply and reapply the terms functioning as observables. The question is, rather, as to the particular choice of terms to utilize as observables.

---

[18] I discuss the problems of operational definitions in much greater detail in "The Logical Effects of Theory on Observational Categories," chap. 3; Also see Hugh G. Petrie, "A Dogma of Operationalism in the Social Sciences," *Philosophy of the Social Sciences* 1 (1971): 145-60.

But now the other line of argument mentioned above concerning our inability to find a unique set of observables comes into play. I have argued that the observation terms in an operational definition serve the logical role of picking out those categories whose criteria of application we assume give us no trouble. But if we cannot locate or describe any absolute or independent set of these categories which could serve as ultimately observable for all operational definitions, then we are forced to admit a relativity of operational definitions. What one theorist treats as an observation term for his operational definitions, another might believe needs itself to be operationally defined, in which case it would play the role of a theoretical term. One man's reliability problem may be another's validity problem.[19] In short, the theory dependency of observation infects operational definitions themselves, with no obvious way to settle on some single set of observation terms. It is indeed a nice philosophical problem as to just how, if at all, one *can* justifiably settle on a set of observation terms for any given investigation. I have argued elsewhere that at least a part of the answer is that we must accept the well-confirmed empirical laws of the moment even though from a logical point of view they may be systematically misleading. That is, we accept as observation terms only those concepts which we have no serious theoretical or empirical reason to believe unreliable. However, such concepts may well prove not to be descriptive of the world. For example, people spoke (and we still do) of seeing the sun rising in the east, even though the horizon is turning away from us.[20]

The application to actual practice of this epistemological doctrine of the relativity of observation to theory is so simple and so widespread as to be almost unnoticed because of its ubiquity. The use of an ordinary control group (plus, of course, extraordinary control groups) is an example. Let us clear our minds of all the bias and prejudice and methodology we have learned and reconsider the situation. A psychologist wants to find out if a certain treatment has a certain effect. He tests, administers the treatment, and tests again to find out if the treatment had any effect. (One could really get down to bare bones by omitting the pretest, which is itself a kind of control group, but the obvious possibility of already existing capacity prior to the treatment makes this supposition strain credulity a bit too far.) On this model we see the pretest as simply a measure of current competence. However, if our "theory" (it need not be terribly precise) enables us to see the pretest also as an opportunity for practice, then it might well occur to us that the resulting effects could be as easily due to the practice as to the treatment. Hence, with a creative methodological leap someone might, looking at the situation in this way, design the experiment with a control group to whom the pretest and posttest are administered but for whom the treatment is omitted. For the experimental group, treatment is included and the results of the two groups compared, and if a difference is noted, it is ascribed to the treatment.

---

[19] See, for example, D. T. Campbell and D. W. Fiske, "Convergent and Discriminant Validation by the Multitrait-Multimethod Matrix," *Psychological Bulletin* 56 (1959): 81-105.

[20] Hugh G. Petrie, "The Strategy Sense of 'Methodology'," *Philosophy of Science* 35 (1968): 248-57.

It would be well to point out here that, commonsensically, there is nothing more "observable" about a pretest than practice. They are just different descriptions of the same event. Some, under the effects of the long-standing positivist mythology, might object that the pretest must be inferred to be practice as a result of the control group experiment. Unless it actually has an effect on the result it cannot be counted as practice and hence is not observable *as* practice. There are two responses to this objection. First, it is a piece of ad hoc linguistic legislation, for it follows from this view that one could never discover that practice did *not* have an effect on later performance since practice is defined as precisely that which does have such an effect. Of even more importance, however, is that this just pushes one step back the ultimate decision as to what shall play the role of observation and what the role of inference. After all, it might equally be argued that the category of pretest must itself be inferred. For it will be a pretest only if we can reliably establish that the subjects are actually displaying their competence and not playing a game with the experimenter which consists of giving as many wrong answers as they can.

Nor will it do to treat the above discussion as a reason for adopting a most extreme form of behaviorism. For even if we eliminate such mentalistic notions as "playing a game," the behaviorist too must establish the reliability of his operational definitions in terms of, for example, gross muscular movements. Even here, mistake is possible and confounded results can occur. The point is simple. At *any* level a question can be raised about the reliability of the application of an observational term, and in the absence of a unique specifiable observation language, the terms which *are* treated as observational will depend on the implicit or explicit theory.[21]

The problem with the foregoing example is that the theory (or common sense) which determines the use of control groups is by now so well confirmed and widely accepted that it may be more than a little difficult for people to be convinced that someone might want to fiddle with its observation terms. For this reason let me give another example. Even with the use of a control group of one form or another, the possibility remains that the attribution of the result to the treatment, though dictated by the logic of the experimental methodology, might yet be in error. One of the most striking recent examples of this is the work of Rosenthal.[22] Roughly the situation is this: Rosenthal has demonstrated the widespread effect of experimenter expectation and the effect of the affective tone of the experimenter on experimental results which had up until that time been wholly ascribed to the experimental treatment. For example, in my terminology the observation term "administration of treatment" was not at all reliably connected with the results. A more appropriate observation term was, e.g., "administration of treatment with certain affective tone." We do not merely see a certain event, we see it under certain aspects.[23] And once again, although the *role* of observation terms must be played by some of the concepts, it is a theoretical-empirical question as to which concepts can best play this role. No unique set of observation terms seems to have any more a priori justification than any other set.

---

[21] The foregoing discussion of the relativity of operational definitions is mirrored in Broudy's discussion in this volume of Polanyi's concepts of tacit and focal knowing.

[22] R. Rosenthal, *Experimenter Effects in Behavioral Research* (New York: Appleton-Century-Crofts, 1966).

[23] Wittgenstein's discussion of "seeing and seeing as" is most illuminating in this regard, *Blue and Brown Books* (New York: Harper Torch Books, 1958); idem, *Philosophical Investigations*, 3d ed. (New York: Macmillan Co., 1968). I have written a paper on this subject which was delivered at the American Philosophical Association, Western Division Convention, May 1970, St. Louis, Missouri. See also my chapter, "Science and Metaphysics: A Wittgensteinian Interpretation," in *Essays on Wittgenstein*, ed. E. D. Klemke (Urbana: University of Illinois, in press).

IV

What I now want to do is to illustrate the thesis of the theory dependency of observation in five selected problem areas—the question of what is learned, the problems of hierarchical structures in learning, the problem of latent learning, the law of effect, and the area of developmental psychology. In each of these areas I will attempt to show how various theories and approaches end up being very nearly incommensurable because of differences in what they allow to be seen. I shall not attempt any detailed exposition of various learning theories because very few are coherent or complete enough to warrant exposition, and because I suspect almost no one would be willing to be identified with such an exposition. I am more concerned with illustrating the theory dependency of observation than I am in evaluating alternative learning theories, and I believe this can most easily be accomplished by concentrating on several issues in detail.

The controversy in psychology over the question "what is learned?" is sometimes presented as a dispute between those who would assert that some sort of central underlying process or disposition is acquired in learning as opposed to those who believe that various pieces of observable behavior are acquired. This is, however, misleading. For at least some of the controversy involves the choice of an appropriate set of observational categories without necessarily involving "underlying" processes or "hypothetical constructs" or "intervening variables" at all. It was pointed out by Campbell that the acceptance of the position that all we have to go on in constructing our psychological theories are the responses of the organism does not thereby commit one to supposing that no central states can be legitimately inferred.[24] Nor does this imply that there is only one way of observationally categorizing the behavior which must serve as our grounding. In short, what is learned may go considerably beyond any simple categorization or combination of observed behaviors.

---

[24] Donald T. Campbell, "Operational Delineation of 'What is Learned' via the Transposition Experiment," *Psychological Review* 61 (1954): 167-74.

Consider one of the experiments discussed by Campbell. In this experiment a conditioned finger movement was obtained through pairing a shock and a tone. The shock could be removed by an extensor movement of the finger. What was learned? An extensor finger movement or withdrawal of the finger? The question is open to experimental investigation. Turn the hand over and repeat the experiment. Now an extensor movement does not remove the shock, but finger withdrawal will. Ninety percent of the subjects withdrew their fingers, 10 percent continued the extensor movement.

It would seem that what was learned by most of the subjects was an action as contrasted with mere behavior. But now many behaviorisms such as Thorndike's connectionism, Guthrie's contiguous conditioning, along with the current eclectic functionalism, would almost all be committed to an explanation in terms of first having learned an extensor movement, next having transfer of training of a muscle group, and a response gradient reaching the new muscle group. Doubtless something like spread of effect would also be invoked. On the other hand, cognitive theories such as Tolman's sign-learning theory or some of the newer information-processing theories would see an action as the observational concept which can be applied to any number of quite dissimilar muscle movements or pieces of mere behavior.

For the associationists the units of behavior tend to be particulate bits of mere behavior and complex behavior must be inferred. And yet in this case their methodological bias seems to require a whole host of strange-sounding entities such as gradients, transfers, and so forth. And even if these "entities" do not have ontological status, nevertheless they stand for very complex processes which seem more and more ad hoc or else seem to require a retreat from a molar behaviorism to physics and physiology in order to be completely clear and unambiguous. The molar criteria for application of the observational categories in the case of the finger withdrawal become tremendously complex, and the predictions which can be made are extremely limited in scope.

The cognitivists, on the other hand, use a set of observational categories in the finger withdrawal case that have as criteria of application some resulting state of affairs which might well be reached in a number of different ways. It would seem far more appropriate with the strictures of a molar behavior theory simply to grant that what is learned is to withdraw one's finger at the tone and that this response is directly observable. The predictive power of this view would seem much better, at least in this case.

Skinner's operant analysis gives us yet a third view of the situation. Skinner's observational categories tend to include as a part of the criteria of application the particular reinforcement schedule applied to the operant behavior, i.e., to the behavior emitted naturally. Given this particular experiment, an operant analysis may not even be possible since the language is that of classical conditioning—usually a much coarser-grained language than is needed for operant conditioning. We simply do not know what behavior was shaped by the particular reinforcement histories. What is learned for Skinner is a sequence of behavior precisely dovetailing into the particular reinforcement schedule which was operative.

Can anything be said about these three different views of what is learned? First of all, no one of them can be said to be "closer to the facts" than any other, for if the preceding discussion is at all correct, the *facts* are different for each of these views at least in the sense of what plays the role of observation and what is inferred. (This is not to say that we cannot make some translations. My descriptions of the situation hopefully provided clues for such translations. It is because I earlier rejected the extreme form of the theory-dependency-of-observation thesis that I can now make sense of some kind of similarity among the positions.) Nor will it do to say that one position makes use of underlying properties while another does not. I have already argued that all are committed to underlying properties which allow for reidentification, albeit the properties differ from theory to theory. According to my previous discussion, the only plausible way of deciding between the theories is by undertaking a thorough theoretical cum empirical examination of their presuppositions, simplicity, predictive power, and so on.

It might be useful to point out here that on the above criterion for evaluation of the various theories, common sense as reflected in ordinary language (the theory of common sense) probably stands head and shoulders above all the other theories. Of the three, cognitive approaches seem best for this particular case insofar as they share, with common sense, observation categories which are of actions rather than of mere movements. Categories applied in terms of intentions and results are often much more reliable than those applied in terms of reinforcement schedules or spatiotemporal bodily movements.[25]

Let me take another example. Consider two baseball pitchers, one left-handed, one right-handed. Suppose further that one has learned the game wholly by playing, reinforced only by natural contingencies; whereas the other has had the benefit of sustained coaching. Now suppose in a game both of them pick a runner off first base. It is hard for me to imagine two more different reinforcement histories and sets of muscular movements, and yet, in terms of action categories, these two men have quite obviously done the same observable thing, picked a runner off first.

It is somewhat surprising to realize that behaviorism, that general program in psychology most concerned about getting as close as possible to the observable facts of the world, uses observation categories which are, in so many cases, so very far from the simplest, most reliable categories they could use—viz., the action categories of ordinary language. The claim that human action is just as observable as human movement is probably a defensible part of the emphasis put on *verstehen* by Max Weber.[26] To observe human behavior with *verstehen* is to see it as action and not mere movement from which action must be inferred.

But if the foregoing constitutes a kind of defense for the use of observational categories of human action in learning theory, at the same time it seems to ignore the 10 percent of the subjects in the finger experiment who continued to utilize extensor movement. *They* did not learn to withdraw their finger. The point to be noted here is that various kinds of conditioning theories do seem to account more adequately for certain kinds of learning than do cognitive theories. We sometimes do learn things "by rote."

The question then arises as to whether or not one needs to invoke a hierarchy of types of behavioral units, or whether these apparently different types of things can be explained in terms of one another.

---

[25] See Krimerman's contribution to this volume for a theory which uses an action-observation language.

[26] Max Weber, *Methodology of the Social Sciences* (Glencoe, Ill.: Free Press, 1949). See chapter by Strike in this volume.

Consider first the attempt to explain away the more cognitive type of behavior in terms of mere muscle movement. Taking a theory such as Hull's, we find that in order to account for what the cognitivist would call insightful or rule-governed behavior, use is made of such concepts as stimulus generalization, fractional antedating goal response, and habit-family hierarchy. The first concerns essentially how any number of physically quite dissimilar stimuli can all be said to be associated with a certain class of responses. For example, think of all the physically dissimilar instructions that might be written, printed, spoken, etc., to get someone to respond. Under a Hullian type of explanation one should be able to account for this by means of a generalization from the originally learned stimulus. The fractional antedating goal response is meant to be used with the habit-family hierarchy to explain why any number of pieces of behavior can be utilized to attain a specific goal. (Compare my example of the baseball pitchers.) The idea is that because of various reinforcement gradients different specific behavior patterns are differentially conditioned to the goal stimulus. Thus under certain conditions one pattern will emerge, while under other conditions another will. This integrated set of patterns constitutes the habit-family hierarchy which is derived from the elementary behavioral units.

Schematically the problem with this kind of handling of the different types of things which are learned is that of specifying the range of the stimulus generalization and the extent of the integration of the habit-family hierarchy. What tends to happen is that either the principle of generalization of the stimulus begs the question by covertly importing the very cognitive term to be defined, or else remains totally unconvincing empirically. In fact, it is only the continual equivocation between sneaking in the terms to be defined and moving back to empirical accounts that has allowed such a notion to survive as long as it has.[27]

---

[27] The schematic strategy outlined above of showing either question-begging or total implausibility is just the strategy used by N. Chomsky in

On the other hand, however, associative learning is not likely to be assimilated to cognitive learning. The lower animals simply do not seem to display the requisite cognitive functions and yet they do learn in apparently associative ways. There is also a philosophical reason for supposing that not all learning, even in humans, makes use of concepts like the gestaltist's "insight" and "organization." Wittgenstein points this out in discussing why reference to obeying a rule is not ultimately explanatory.[28]

Briefly the reason is that rules can always be variously interpreted and until we ground the infinite regress of rule, interpretation of the rule, another rule to enable us to understand the interpretation, yet another interpretation, and so on, in some kind of actual behavior in the world (form of life), we will be unable to explain the correct interpretation of the rule.[29]

What emerges is the necessity for a nonreducible hierarchy with genuinely different criteria of application of the terms at the various levels. What the associationist needs to recognize is that no matter what the fine details of the associative bonds formed between bits of behavior, or no matter what the reinforcement history, the criteria of application of the concept of what is learned (and what will thus be reinforced) do not, in some cases, refer to these bonds or this history. The concept of what is learned is often

---

his celebrated "Review of Skinner's Verbal Behavior," *Language* 35 (1959): 26-58.

Charles Taylor, *The Explanation of Behavior* (New York: Humanities Press, 1964), also mounts a sustained attack on the ability of simple S-R mechanisms to account for complex behavior even in animals. Taylor's major strategy is to show the ad hoc character of the explanations introduced to handle these kinds of embarrassments, to note the experimental implications of these ad hoc explanatory principles, and to point out that where such experiments have been carried out, the ad hoc principles fail in the same kind of way. Since any theory can, logically, be saved by adding epicycle on epicycle, only a thorough critique of this kind can ever discredit a theory. And even then, as Kuhn has pointed out, the discrediting will be effective only when a powerful alternative theory is available. See Thomas Kuhn, *The Structure of Scientific Revolutions* (Chicago: University of Chicago Press, 1962).

28. Wittgenstein, *Philosophical Investigations*, sec. 85.

29. See my detailed discussion of the philosophical issues involved here in "Science and Metaphysics: A Wittgensteinian Interpretation," in E. D. Klemke, *Essays on Wittgenstein* (note 23).

analyzed in other ways, such as in terms of intentions, or actions, or results of the behavior, independently of its fine details. But such concepts can themselves become the units of behavior, i.e., the observables, for a new round of learning. Concomitantly, the gestaltist or cognitivist must realize that to characterize all learned behavior with such observation terms seems to ignore the 10 percent who learned the extensor finger movement. But even more importantly, it gives up the explanatory value of classifying some learning as insightful. For such a classification is explanatory only as long as it is used to mark off a different set of observation categories from the ones traditionally used by the associationists. A part of correct explanation is good classification as long as the classification marks a real difference.

It is no accident that information-processing theories of learning sound very "cognitive" in their language. This is at least partly because these theories have incorporated the notion of a hierarchy from the beginning. They take seriously the fact that the higher-order processes not only can represent lower-order processes in one way or another, but also that these higher-order processes have other features of their own. Just as a mathematical theory may represent reality, there are syntactic and purely mathematical features of the theory which are incidental to its representing function but nevertheless quite real. Thus the categories in the higher levels may represent the lower levels, but not be identifiable by means of reference to the lower levels. The hierarchy is nonreducible. On the other hand, however, information-processing theories are just as "mechanical," "hardheaded," and "nonmysterious" as the most dedicated behaviorist could wish. The continuing success and fruitfulness of computer simulation techniques utilizing such theories have graphically shown that low-order mechanical reactions can be so organized as to produce highly organized "meaningful" behavior. In short, simulation has shown the associationist he need not fear mentalistic vocabularies and has shown the cognitivist that his higher-order concepts can often be realized in a totally nonmysterious way.

The point is that if we attempt to describe all experience in associationist categories, certain features, e.g., the higher-order functions, seem to get left out. Associationist theory simply seems to have no room for a hierarchy even if the hierarchy is ontologically unobjectionable, as in information-processing theories. Cognitivist theories, however, seem to remain suspended in the air when an attempt is made to apply their observation categories to everything. Both associationists and cognitivists must be brought to realize the blinders which their theories impose on their observational categories.

I turn now to the phenomenon known as latent learning. This refers to the fact that in a variety of situations behavior goes on which does not at that time seem to result in any learning, i.e., no activity is originated or changed at the time, but is such that at some later time the learning is elicited. For example, rats able to explore a maze without reward begin, as soon as they are rewarded, performing as well as rats trained all along with a reward. A more homey example is the student studying for a test who does not perform until the test occurs.

This phenomenon is often marked by transformational linguists as the distinction between performance and competence. Although *some* kind of performance must usually be used to determine if competence has been gained, this need not always involve any straightforward exercise of the competence. For example, we might simply accept the performance of an actor who is trustworthy and who claims to have memorized his lines.

Although competence is probably not one of the observation categories of even the transformational linguist (i.e., it is inferred for him, too), it is much easier for him to handle since he does not operationally define competence in any single-track sort of way. Given that certain kinds of adaptive actions *are* observable, he simply postulates the competence as one of the hierarchical entities which help give rise to the actions. The competence has a life of its own independent of the actions which occasionally exemplify it. The justification for such status is the theoretical integration given to the actions supposed to flow from a certain competence.

A similar move is not nearly so easy for associationists to make. In the first place, they tend to define completely their inferred terms by means of operational definitions. This then cuts off any independent life for these terms. The inferred term *is* the operational definition and it makes no *sense* to ask if there are other operational definitions of the inferred term or whether the original definition should be modified. Such questions presuppose that the inferred term has some use beyond its operational definition. But even if the associationists allowed themselves a bit more freedom in the ontological status of their inferred entities, they would still face problems in accounting for latent learning because of the limitations of their observation categories. Let us see why.

Many associationists adopt a drive-reduction view of reinforcement. That is, crudely, they have observational reasons to believe in certain kinds of physiological and biological drives, e.g., hunger, thirst, and so on. Reinforcement occurs when one of these drives is reduced. Thus the drive causes behavior. When a certain pattern of behavior succeeds in reducing the drive, that pattern is reinforced and learning has taken place. Now the problem begins when the independently identifiable physiological and biological drives seem unconnected with the phenomenon of latent learning. In fact, that is just another way of specifying the area of latent learning. It is behavior continued and learned without the customary drive reduction. Typically, the associationist postulates new drives, as, for example, curiosity. The drive-reduction theory of reinforcement remains in effect and all is well. But not really. For given the behavioral observational categories of the associationists, the *only* indication of the new drives postulated to account for latent learning is the very behavior which manifests latent learning. There is no independent way to specify the drive. The assertion that there is a law-like connection between the drive and behavior which reduces the drive becomes tautological. With the limited observation categories of the associationist, there seems no plausible way to avoid such a result.

Notice here, however, that Skinner's notion of operant behavior seems to avoid the problem. For Skinner's basic observational category of operant behavior does not require him to see the behavior as caused. Where other associationists see two events, Skinner sees but one.[30] Thus Skinner is obligated to explain the change of behavior but not the origination of behavior. For *what* is observed is not just behavior, but behavior of a certain kind, i.e., of the kind connected with the particular reinforcement schedule.

---

[30] James E. McClellan, "B. F. Skinner's Philosophy of Human Nature: A Sympathetic Criticism," in *Psychological Concepts in Education*, ed. B. Paul Komisar and C. B. J. MacMillan (Chicago: Rand McNally & Co., 1967) makes a similar point.

The cognitivist also avoids the problem by not observing merely "action," but rather action of a certain kind. His criteria of identification refer to results and intentions rather than to reinforcement schedules, but in both cases a level higher up in the hierarchy is treated as containing the observation terms, thereby making inferences to higher levels a bit more plausible. In both cases, however, the directions which the theories can take in meeting their problems and the limitations they encounter seem closely bound up with what the theory says is observable.

The apparent circularity of the explanation of latent learning by the drive-reduction theory of reinforcement is similar to the oft-repeated charge that one of the major laws of learning, the so-called law of effect, is also circular. Although the law of effect is stated in innumerably different ways, examination of just one of these should suffice to show the possible pitfalls. Consider the following ". . . a learnable response followed by a reinforcing event (stimulus, state of affairs) will receive an increment in its strength or probability of occurrence."[31] The charge is then made that unless we can independently specify what a reinforcing event is, the law of effect becomes a definition of a reinforcing event and hence circular. This charge, however, generally does not pay sufficient attention to the subtlety with which associationists deploy the law of effect. In the first place, the law will be truly circular only in case there is not *any* independent description of the reinforcing event other than that it increases the strength of the learnable response. Of course, most of the time there is. What this exemplifies is the general principle that logical relations such as tautologicality depend on the descriptions used rather than on the referents of these descriptions.

Suppose, for example, that "*a* caused *b*" is a true singular causal statement—a paradigm case of an empirical statement. But, ex hypothesi, "$a =$ the cause of *b*," and by substitution we get "the cause of *b* caused *b*," an analytic statement in which the description of the cause is *not* independent of the description of the effect. Yet one is under no compulsion to assert that the connection is not empirical.[32] The principle is generalizable.

---

[31] Hilgard and Bower, *Theories of Learning*, p. 482.

Thus to characterize an event as a reinforcer is at least partially to claim that it is observably identifiable in terms of its effect on certain kinds of learnable responses. But that is *not* to say that the reinforcer may not be identifiable in some other way as well. Under this view it is possible that a reinforcing event may not reinforce. And there is no paradox here. The situation is similar to the lack of paradox in saying, "Jim was shooting baskets yesterday, but he didn't shoot (make) a single one." In both cases the activity is identified partially in terms of what we know generally are the results of, or characterize, the activity. We are using observation terms which generally have a fairly wide range of application in the sense of classifying many otherwise dissimilar events as events of the same kind. The law of effect thus provides a learning-theoretic set of observation categories for behavioral events that can also be described in other terms which could in turn show the law of effect in particular cases to be empirically false.[33]

Such a view becomes even more plausible if we note that whether a reinforcer really reinforces depends on other things than the fact that it usually does so. It might also depend on the subject's noticing it, for example. The situation is analogous to the fact that a perceptible object will sometimes not be seen, even though it is in the subject's field of vision, unless other conditions are right, e.g., the subject is paying attention. Once more the influence of theory on observation terms is apparent.

All the preceding examples do exemplify, I believe, the relativity of observed and inferred and the theory dependency of observation. Nevertheless, it is probably the case that because many readers will themselves be committed to a particular theory, the phenomena I have taken to illustrate the theory-dependency-of-observation thesis will be seen by them as simply my pushing a particular substantive point of view (despite the fact that I have tried to accuse nearly everyone of parochial observation at one point or another). Part of the cure for this simply involves going over the material again and again and trying to take the other person's viewpoint. If you are an associationist, try to see (not infer) actions instead of movements in some instances. If you are a Skinnerian, try to see behavior through observational categories that do not have the particular history of reinforcement as part of the criteria of application. If you are a cognitivist, try to see that certain rudimentary behaviors need to be learned associatively to provide the basis for the application of your own (differing) observational categories. If you are a critic of associationism, try to see certain events *as* reinforcers rather than seeing one event, then another event, and inferring reinforcement.

---

[32] This example is drawn from Donald Davidson, "Actions, Reasons, and Causes," *Journal of Philosophy* 55 (1963): 685-700.

[33] I believe this characterization of the law of effect as providing a new set of observation categories is compatible with that given in P. E. Meehl, "On the Circularity of the Law of Effect," *Psychological Bulletin* 47 (1950): 52-75.

In addition to such reiterations of what has gone before, let me offer one more example of the influence of the theory dependency of observation. This example is from developmental psychology, and because of the way it involves our most philosophically basic observational categories, it may be somewhat more persuasive than the preceding.[34]

If, following Kessen, we take cognitive psychology to be concerned with reality and man's representation of reality and if we attempt to explain how a child can grow up to master mathematics, physics, psychology, and even philosophy, it is clear that our analysis must necessarily be constrained by our own view of the world.

The problem in a nutshell is that we as adults define the terms and set the developmental goals, in terms of which we describe and evaluate a child's progress. It is the reality as adults structure it that we foist off on the child. There are some nice philosophical puzzles here about whether we can logically avoid such an imposition and to what extent we are justified in believing that our own adult conceptual scheme adequately represents reality, but although such questions are clearly relevant, they have often been discussed before. What I want to do here, however, is to assume that somehow, perhaps evolutionarily, the adult conceptual scheme is adaptive to, if not closely representative of, reality.[35]

---

[34] See, for example, William Kessen, "Questions for Theory of Cognitive Development," *Child Development Monograph, Serial* 107, 1966, Vol. 31, No. 5, pp. 55-70, for a psychologist's view of such problems.

With this assumption in hand I want to call attention to two current controversies that fit into the problem structure of developmental psychology. I refer to the revival of the nature-nurture controversy by the ethologists[36] and to the women's liberation movement.

Turning first to the ethologists, quite clearly one of the major problems is whether we are to see certain kinds of human behavior as essentially aggressive, or are we to see aggression as aberrant and caused by some lack of genetic determination or upbringing. Or is it some combination of the two? I do not wish for a moment to endorse what I find to be oversimplistic global explanations by the ethologists. What I do wish to do, however, is to point out that at least a part of what they are asking us to do is to *look* at the world in a different way—a way conditioned by their theories. The questions they raise are as much about the categories of observation as about theoretical interpretations. To give just one example—if aggressive behavior is "natural," then aggression can be counted as one of the basic physiological drives and used in explanations of other aggression-related behavior. If aggression is not natural, then *it* must be explained in some way by other drives or stimuli which are themselves assumed to be natural.[37]

The case of women's liberation is even more fascinating. Discounting the more rabid proponents, there is a most intriguing picture presented by the movement. Essentially, they are saying that the categories of femininity presented by the dominant male culture are *not* based on basic biological or physiological characterisitcs, but are rather the result of the cultural picture the adult male has of what a truly feminine human being should be like. The pervasive character of the picture is explained by the fact that women become "willing, unknowing participants" in their own subjugation. One of the interesting points here is that even after one attempts to determine empirically what attributes can be attributed to upbringing, the movement continues, at least in some instances, to make out a plausible case that the very categories in which the experiments are carried on are themselves male-culture–bound. For example, it is at least plausible to entertain the hypothesis that some of the physiological differences one finds between male and female and which show less strength on the part of the female are traceable to differences in the amount and kind of physical exercise our culture deems important for the two sexes.

---

[35] I discuss this problem in more detail in "Science and Metaphysics: A Wittgensteinian Interpretation," see note 23.

[36] E.g., Konrad Lorenz, *On Aggression* (New York: Harcourt, Brace & World, 1966); Robert Ardrey, *The Territorial Imperative* (New York: Atheneum, 1966); Desmond Morris, *The Naked Ape* (New York: McGraw-Hill Book Co., 1968).

[37] There is a close connection between what is counted as observable and what must be inferred and what is natural and what must be explained. See Toulmin, *Foresight and Understanding* (note 4).

Notice that I am not talking about the fairly obvious sociocultural differences in treatment of male and female, but rather about the basic categories of masculinity and femininity themselves. Do we start by seeing human beings and determining what differences there are, or do we start by seeing males and females and determining what similarities there are?

The analogy with child development is a good one. We cannot simply *use* without examination our basic categories of time, space, causality, etc., in describing the attempts of the child to construct these very categories. If we do, somehow we will likely miss the most important features of the process of construction; namely, are there reasons, other than the fact that adults reinforce their use, for developing the categories we do. The case with masculinity and femininity is similar. We cannot just use, without explicit awareness, attention, and justification, the present categories of masculinity and femininity to describe the attempts of women's liberation to question these very categories. Indeed, we must start somewhere, but the women's liberation movement has shifted what they will count as observable and what they believe must be inferred, and for us stubbornly and dogmatically to insist on a different set of observation categories even backed up by a simple elaboration of the theory which determines *our* categories is to miss the point entirely. Once more the relativity of observable versus inferred and the theory dependency of observational categories shows us why there is so much heat and so little light. The opponents simply talk past one another.

V

What methodological morals can be drawn from all this? For one thing, I believe it shows why "learning" as used in scientific theories of behavior such as Hilgard's bears so little relation to "learning" as used in the classroom. It is not that the behavior-theory concept of learning is just more refined or broken into smaller parts than the classroom concept. It is rather that the rough-and-ready theory of the classroom uses blatantly mentalistic action categories as observables, whereas there is a great deal of resistance to this in behavior theory. Indeed, the very name "behavior" betrays the tendency not to use action categories.

Further, I have tried to point out various implications for research practice as I went along. I have urged careful and continuing attention to the correct description of "what is learned." I have suggested that more consideration be given to theories which claim that human action (intentional behavior) is directly observable. I have pointed out that most behaviorisms, rather than getting closer to the facts in the sense of getting closer to that which can justify our explanations and theories, perversely seem to use categories farther away from the highly reliable ones of common sense. I have urged the necessity for recognizing hierarchies in accounting for learning. I have noted the difficulties surrounding the definitional and empirical use of some of the key terms in learning theory. In this regard I have urged that "reinforcer" as used in the law of effect can have the role of an observable. Finally, I have claimed, especially in the areas of developmental psychology, that the philosophical presuppositions and justifications for the categories in use must surely be displayed in those cases where these categories are themselves being questioned.

In addition to these methodological strictures, let me add a few remarks of a more general nature. More importantly, researchers must stop wasting their time looking for crucial experiments. It should be obvious by now that any theory can be saved from embarrassing empirical consequences if we only add enough "epicycles." But, more profoundly, the crucial experiments tend to be crucial only if conducted within a single research paradigm—that is, only if they operate from the same theory and hence with the same observational categories. For without these common denominators, two theories will just pass each other by through a failure to countenance the same facts.

A corollary of the abandonment of the crucial experiment is the necessity for empirical researchers constantly to be theorizing as well and to be elaborating their theory with its presuppositions as they go. Even if one could argue that this constant appeal to fundamentals is wasteful, at a minimum one must constantly remember that the fundamentals *do* infect ordinary research and be a bit more tolerant of those whose concern is for fundamentals. For the only ultimate way to justify a position and defend it against plausible rivals is by means of a detailed theoretical cum empirical cum philosophical exposition of its own principles and a similarly detailed polemic against the rival positions *in their own terms.* In Kuhn's terminology, only well articulated paradigms can be confirmed or refuted.[38]

Finally, the foregoing gives some support to the thesis that the dominant position in current American psychology—an eclectic functionalism—must be rejected. What I am referring to is a loose attitude of tolerance and borrowing from many traditions while working in a hardheaded way on small localizable problems. Borrowing from many theories without carefully evaluating and incorporating the theory which renders intelligible that which is borrowed can only distort and falsify. The inability to see that some empirical results have the significance they do because of the observational categories and theory in which they are embedded is functionalism's greatest sin. Indeed, the "tolerance" of varying viewpoints seems attractive and fair. But this tolerance is a primrose path shutting out the influence of theory on observation and leading to the stagnation of psychological research. Theorizing is no sin. It cannot be avoided in any event, and claiming that psychology is too immature for theory construction ensures its fixation at that immature level.

---

[38] Kuhn, *Structure of Scientific Revolutions.*

SECTION II

CRITICIZING CURRENT CONCEPTS

CHAPTER IV

Wanted: New Paradigms and a Normative Base for Research

HAROLD B. DUNKEL

The most fruitful area of convergence between philosophy and educational research has been that which can be vaguely labeled *science*—as many of the chapters in this yearbook show. Since the startling advances and rising prestige of science have made it the area where the action is, it is scarcely surprising that workers in adjacent fields have moved over as close as possible in order to share the aura. Consequently, what is commonly known as philosophy of science has probably been the field of philosophy most worked in recent decades; and from the other side, educational psychologists, sociologists, and economists (like the practitioners of their respective parent disciplines) have made intensive efforts to be "scientific." The consequences of this movement from both sides has been a common interest in terms and definitions, scientific paradigms and experimental designs, inference (particularly the logic of necessary and sufficient conditions), and the like.

Education both as a field of study and as an activity has benefited much from this collaboration. Yet, an effort by some educators to be more scientific than the Nobel laureates has had one unfortunate consequence—a tendency to define *science* and *scientific* in an extremely limited fashion. A roll call of the sciences, mustering them under the various sets of familiar rubrics such as the "hard" and the "soft," the "pure" and the "applied," and the "natural," the "social," and the "humanistic" reminds us that the term *science* can comprehend a very wide variety of subject matters, methods, conceptual structures, and purposes. (We do well to recall one of Max Weber's examples of "a scientist" in *Wissenschaft als Beruf*: the philologist who says this emendation is to be made in this manuscript at this point.[1]) In spite of this heterogeneity among the activities which can properly be labeled science according to some definitions of it, many philosophers and educators have customarily constricted the range of application very tightly. To be sure, most formal statements concerning the nature of educational research exhibit a more catholic attitude, although, for example, Kerlinger in the *Encyclopedia of Educational Research*[2] lays primary stress on the points characteristic of what I am calling the narrow view. Then, too, there are probably the usual differences between what professors say, what they do, and what impressions both of these activities leave with their students. Evidence of strict constructionism is fairly apparent in the concept of "real research" traveling on the graduate-students' grapevine in most departments or schools of education. Certainly not all philosophers and not all educators have construed science so narrowly; and those who do so are, of course, entitled to their definitions. Nevertheless, this fairly general trend has consequences for educational research and, ultimately, the educational process.

Philosophers, insofar as they have concentrated on science in this restricted sense, have offered fewer suggestions and analyses in other areas. Workers in educational research, on their side, have been rather uninterested in the "nonscientific" parts of philosophy, and in their own work they have tended not to undertake the sort of inquiry which would arouse the interest of "nonscientific" philosophers. The purpose of the present chapter is to indicate a few gaps between the groups that need bridging or at least to suggest places where a tree or plank already spanning the chasm might well be replaced by some more spacious and reliable structures.[3]

This narrow view of science tends to define it by three characteristics: (1) as descriptive, (2) as value-free, and (3) as nomothetic. Science as thus conceived has limited utility for education, great as the utility within those limits may be.

---

[1] Translated in H. M. Gerth and C. W. Mills, *From Max Weber: Essays in Sociology* (New York: Oxford University Press, 1946), p. 135.

[2] F. N. Kerlinger, "Research in Education," in *Encyclopedia of Educational Research*, ed. R. L. Ebel (New York: Macmillan Co., 1969), pp. 1127-41.

[3] Twenty years ago G. E. Barton, Jr., in a very different context suggested a bridge-building operation. ("Educational Objectives—Improvement of Curricular Theory About Their Determination," in *Toward Improved Curricular Theory*, ed. V. E. Herrick and R. W. Tyler [Chicago: University of Chicago Press, 1950], pp. 26-35.) Although his article is still cited in the bibliographies, it seems to have produced little construction.

(1) According to this definition science is purely descriptive, the investigation of the relations among phenomena. But education is not, strictly speaking, descriptive although, of course, it makes great use of descriptive knowledge. Education's primary phenomena would simply not exist if it did not produce them itself. It is an "operative science" (to use Bacon's term) or a "productive science" (to use Aristotle's), not a descriptive one. It is an intentional, artificial intervention in the natural development or maturation of human beings.

(2) This intervention through education must be justified. This fact stands in sharp conflict with the definition of science as value-free. The scientist can and does legitimately intervene and alter a natural course of events simply in order to observe or predict the new state of affairs following his intervention. Education, however, is intended to be "good for" students. Those educational agents intervening (parents, teachers, the political state) always intend that the student should be improved in some fashion, however stupid, silly, or perverted their criteria may be. (That the recipient of such attention or that an impartial observer of this educational attempt may deem it a waste of time or a positive bane does not alter this character of the educator's intention. Even those instances where the educational process seems to run on merely from habit, tradition, or bureaucratic inertia indicate only that the original justifications still go unchallenged.) Since, then, education does not aim at mere "change" but at "improvement" of some sort, education is inevitably involved in normative discourse and problems of value to an extent that the "hard" sciences rarely are. (The difference between physiology or pharmacology on the one hand and medicine on the other suggests the same point.) Philosophers who have apparently come close to suggesting that the educator should abjure all value statements seem to undertake the purification of his language at the cost of rendering him mute.

(3) Finally, says the narrow view, science is concerned with universals, or, in more current terminology, with generalizations of extremely high probability. Education, too, wishes to know such

things as what happens in most cases or what is true of most students, as is evident from the perennial interest in norms, standards, and surveys. But this very interest perhaps unduly distracts attention from the degree to which education is idiosyncratic as well as nomothetic. Teachers rarely feel they are facing merely 3-300 incarnations of points on a distribution; they hope they are educating Johnny Johnson and Suzy Smith. But, by those espousing the narrow definition, dealing with the individual is usually considered an affair of art (medicine curing this patient) or technology (engineering building this bridge); the whole conceptual apparatus of science, along with its counterparts in educational philosophy and educational research, is often seen as inapplicable.

In sum, then, insofar as science is characterized as descriptive, value-free, and nomothetic, especially as a combination of all three, it and the developments associated with it and growing out of it leave largely untouched some areas of major concern to education.[4]

Intervention raises the question of the purpose of the intervention. What is intended? What is to be accomplished? Knowing what one is trying to do seems a minimal requirement for rational behavior; and certainly one obvious way of evaluating educational programs is to ask what they were trying to accomplish and then to see whether this end was reached. As a result we have had for the past forty years and more great emphasis on stating objectives in behavioral terms.

Yet a number of things have been trying to tell us that this pattern, this Aristotelian dominance of the final cause, has certain serious shortcomings, notwithstanding its admitted utility on many occasions. For one thing, in spite of the admonitions of superintendents, curriculum consultants, and evaluators, many (probably most) school systems and institutions of higher education have never developed specific statements of their objectives; and where such statements have been produced they have usually been mere paper creations with little relation to actual educational operations.[5]

---

[4] Additional characteristics of this narrow interpretation of "science" could be cited—e.g., that "science" must involve the statistical manipulation of quantified data. But the three points mentioned in the text will suffice for present purposes.

[5] M. Ammons, "An Empirical Study of Process and Product in Curriculum Development," *Journal of Educational Research* 57 (1964): 451-57.

For another, teachers and observers of teachers often insist that they pay little attention to stated objectives even if they have written them.[6] Likewise, more technical analysis of the problems has indicated that there are educational undertakings where this procedure is inappropriate—for example, in creative work in the arts.[7]

Since objectives will be treated elsewhere in this volume, they will not be discussed at length here. But these difficulties do give occasion for surprise that paradigms other than this emphasis on ends and final causes have not received greater attention, particularly since there are a number of contemporary forces which tend to make this traditional model even less acceptable.

In connection with these pressures, two characteristics of stated objectives are especially relevant: (a) they are future-oriented; (b) they concern specific elements in isolation rather than within the total gestalt they form.

Formerly the little boy of four or five, when asked what he wanted to be when he grew up, would reply, "A fireman" or "A doctor" or something of the sort. Nowadays, if he reacts as do many young people fifteen or twenty years older than he, he will respond, "Me." The whole idea of future career slots (with its prerequisites for the prerequisites) or, in general, of some future state to be worked toward as distinct from life as lived now is singularly out of fashion. Teachers and institutions at all educational levels are being asked to talk about education in terms of what it means now rather than in terms of what it leads to. The "free" universities, student-organized and student-directed classes, and similar phenomena on the contemporary educational scene are in large part species of a general revolt against future goals ("mortgaging the living present for the uncertain future") and deferred realizations. To determine how long and how strongly these current forces will persist requires a less clouded crystal ball than is presently available. But research in such fields as curriculum, teaching, and evaluation will at best be "unresponsive" if stating future objectives is the only prevailing mode of procedure.

---

[6] See P. W. Jackson, *Life in Classrooms* (New York: Holt, Rinehart & Winston, 1968), pp. 164-68.

[7] See E. W. Eisner, "Educational Objectives: Help or Hindrance?" *School Review* 75 (1967): 250-60.

The second difficulty, the failure to deal with total configurations rather than isolated objectives, can be documented by well-worn examples. As has often been pointed out, the superintendent or principal who put into his school all the "new" programs (the "new math," etc.) might have a set of good courses, but not necessarily a good curriculum; he would have an agglomeration, not a structure, for these programs were usually evolved by examining the structure of a given discipline, without regard for its relations to anything else, including the student.[8] The curriculum thus becomes a thing of rags and patches rather than a seamless cloak.

We have children, almost from the time they enter the first grade, collecting, dissecting and experimenting with animals, in order to teach them science. But only occasionally does an educational philosopher suggest that we ask ourselves whether we are aware and approve of the attitudes toward living beings and the natural environment we are possibly engendering by these familiar procedures.[9]

Or to take another example, in our enthusiasm over the effectiveness of immediate reinforcement, we have children working for teacher approval, for tokens, for dimes, and for whatnot. Attention is good; interest is high; the stated objectives are achieved as never before. But fascinated by these very real gains, we are possibly failing to worry sufficiently about other objectives we are probably also achieving: teaching the student that the first question to ask in every learning situation is, "What's in it for me?" suggesting to the pupil that immediate gratification is preferable to deferred satisfaction, making him an other-directed addict to immediate external feedback and the like. Perhaps these unstated objectives are desirable, perhaps not. The point is that we often fail to see the issue because of our myopic concentration on a few stated objectives, despite all the warnings that W. T. Harris, John Dewey, and a host of others enunciated long ago about such collateral or incidental learnings.

---

[8] See John I. Goodlad, Renata von Soephasius, and M. Frances Klein, *The Changing School Curriculum* (New York: Fund for the Advancement of Education, 1966), pp. 92-95, 106-7.

[9] J. R. Bronars, "Tampering with Nature in Elementary School Science," *Educational Forum* 33 (1968): 71-75.

The feeling of disquiet can perhaps be made most clear by initially considering the case of classical psychoanalysis. The therapist does not attempt to move the analysand toward some future norm of "mental health," either in general or as a set of particulars. He makes a diagnosis of what is wrong with the patient's *present* pattern of behavior and tries to see whether a better mix can be achieved, even though considerable pathology may be latent within the new pattern simply because the patient can handle and integrate only so much. To be sure, the analyst's thoughts *can* often be translated into the language of objectives and we can make him say: "This patient should be better able to sleep" or "to work" or "to love." But this translation seems to distort to some degree what is going on. The germ of the point is inherent in the stock response of a famous analyst when twitted about one of his patients who was a fairly famous but not very admirable person: "I was not his creator; I was only his analyst." The clear-cut set of intentions may be within the power of the creator or, at least, the artisan; but the analyst and the teacher usually feel no such God-like strength.

In fact, one might almost conclude that educators are driven by a kind of Napoleonic complex when they talk in the military jargon of "classroom strategies," "curricular tactics," "reaching objectives," and the like. Some of us wonder how far such talk is applicable. For educational operations are often more like those military operations on Dover Beach where ignorant armies clash by night than they resemble the thrusts of armored columns (prettily portrayed on the military maps by those long, sweeping black arrows). If we must use military language at all, the teacher seems better characterized as conducting continuous patrol actions (not charges), feeling out spots where he can advance and trying to hold such ground as he can gain; but his "objective" tends to be whatever ground that is. And if he cannot advance as he would wish, he tends to put his hopes in the doctrine of the English carnival proprietor, "What you loses on the swings, you makes up on the roundabouts."

The student confronted by a set of formal objectives is likely to feel he is being thrust into a rat race over a long series of hurdles set before some predetermined goal. Many of the currently popular critiques of education gain their force from their emphasis on the adequate assessment and manipulation of current patterns rather than straining for future goals.

The research area in which the gestaltist or molar approach has usually been most at home has been personality assessment, a consequence of the position taken in the original studies by the staff of the Office of Strategic Services (OSS).[10] But even here disagreements among judges and similar difficulties have tended to drive recent research more toward the scientific—that is, toward the better isolation and more precise measurement of specific traits, with any attempt at combining them in a total assessment falling back on some empirical strategy (e.g., comparing the traits of an individual studied with those of some group deemed typical incumbents of the post for which he is being selected) or on the statistical strategy of factor analysis. But these attempts to dodge the normative problem have all the familiar drawbacks—they perpetuate the status quo rather than improve the breed, they offer no explanation or justification of the pattern as such, and so on.[11]

Considerations like the foregoing suggest the desirability of some paradigms which enable us to deal with configurations, not merely configurations of objectives for future attainment (though such models would be useful even to those who prefer the language of objectives), but existing configurations as well.

The traditional means of judging patterns as patterns have been the age-old aesthetic criteria—unity amid variety, balance, and all the rest. But although Whitehead showed how they could be given much wider applicability and although the gestalt psychologists in their work on "good figure" made one specific application to the field of perception familiar to educators, this procedure has been largely ignored as nonscientific.[12] Yet it remains the most

[10] The OSS Assessment Staff, *Assessment of Man* (New York: Rinehart & Co., 1943).

[11] E. L. Kelly, *Assessment of Human Characteristics* (Belmont, Cal.: Brooks Cole Publishing Co., 1967), pp. 71-86.

[12] A. N. Whitehead, *Adventures of Ideas* (New York: Macmillan Co., 1936), pp. 309-81; H. B. Dunkel, *Whitehead on Education* (Columbus: Ohio State University Press, 1965), pp. 59-73.

established way of assessing goodness of pattern. New criteria are not likely to be discovered at this date. But better understanding of the ways in which the familiar criteria could be applied to educational materials and a vast improvement in the techniques for so doing could be forthcoming.

The possible utility of a different present-oriented, holistic model receives additional reenforcement from the current pressure for activism on the part of both students and teachers. They do not want merely to know what pollution is, how it comes about, or how much there is of it; they want to do something about it. Granted that this last activity necessarily involves knowledge about the preceding questions, certainly the answers to them alone are not sufficient to produce rational, effective action. Visitors to faculty meetings are often surprised that men so skillful in various complex theoretical inquiries prove so inept in deliberation and council. Education in general and educational thought and research in particular have paid less attention to inquiries involving decision and action than to those concerned with knowing. One need not adopt the view that educational institutions should be, or willy-nilly will be, politicized to believe that education in the future will be and should be more concerned with these matters than it has been in the past. If that overused word *relevance* has any referents at all, one of them would seem to connect it with educational involvement in choice and action. For example, a recent and much discussed survey of American education has charged it with being "mindless"; but interestingly enough, the remedy there proposed for this malady was not a new conceptual framework or even a new set of objectives, but an operative procedure—the "open" school.[13]

In view of Dewey's traditional association with education, it may seem strange to suggest Dewey's theory of inquiry as a much neglected model which possesses the virtues for present purposes of being present-oriented, action-oriented, and holistic. But such is, I believe, the case.

This neglect of Dewey appears to spring from numerous possible causes, some of which are connected with his being unscientific in the narrow sense of science noted earlier in this chapter.

---

[13] C. E. Silberman, *Crisis in the Classroom: The Remaking of American Education* (New York: Random House, 1970).

And sometimes his theory of inquiry has been regarded merely as a crude, out-of-date statement of the paradigms for scientific inquiry now better formulated by Hempel and others; but this view ignores Dewey's dialectical merging of the theoretical and practical. Actually, those "scientists" who gave Dewey credit for the attempt, even if unsuccessful in their opinion, to make action and valuation scientific were properly put off by what they necessarily regarded as the perverse way in which scientific inquiry in its turn was infected in his theory by considerations arising from the demands of action and valuing. Or sometimes Dewey has been made into a sort of latter-day Aristotle, with his ends-in-view transmogrified into final causes.

But Dewey's ends-in-view are not distant aims, the kind that can be translated into specific educational objectives without doing violence to his thought. This fact is abundantly clear from the impatience of devotees of "specific objectives stated in behavioral terms" with the man who would say only, "The aim of education is growth"—not a very specific or behavioral statement. As part of Dewey's well-known telescoping of the ends-means distinction, ends-in-view are hypotheses concerning likely *means* of solving the problem which has evoked a given inquiry. They arise from the present problem and return to it as solution. And in talking about the "problem" and "inquiry" we must not ignore Dewey's caution that most of his critics who failed to understand him erred by failing to recognize the emphasis he placed on the present problem as the pivot of inquiry and the degree to which he viewed this problem as "particular" and "differential." [14] As he was always careful to point out, the general theory of inquiry is only an abstraction from hosts of specific inquiries concerned with farming, healing, dancing, and valuing. Inquiry is singular. The "facts of the case" are the facts of *this* case; the relevant "ideas" are those relevant to *this* problem *as here formulated* and to the facts of *this* case.

---

[14] J. Dewey, "Experience, Knowledge and Value: A Rejoinder," in *The Philosophy of John Dewey*, 2d ed., P. A. Schilpp (New York: Tudor Publishing Co., 1951), pp. 559-60. See also his *The Quest for Certainty* (New York: Minton Balch & Co., 1929), pp. 123, 228; *Logic: The Theory of Inquiry* (New York: Henry Holt & Co., 1938), pp. 496-97; *Theory of Valuation*, International Encyclopedia of Unified Science, vol. 2, no. 4 (Chicago: University of Chicago Press, 1939), pp. 46-49.

Deweyan inquiry does have some virtues sought here. It is present-oriented: it attacks a real, existential problem; and even as the inquirer moves toward solution through use of the end-in-view, it is with a backward glance at the formulated problem to make sure that the course of action espoused is indeed a solution to it. Inquiry is holistic in that both problem and solution must take account of as wide a range of matters as possible if inquiry is to be competent and its solution satisfactory. It is action-oriented since it deals with an existential situation; inquiry does not deal with states of mind."[15]

Schwab has recently been emphasizing the need to expand educational experience to include the domain of the Aristotelian "practical" and its "deliberation."[16] I suspect that Schwab has chosen to speak in terms of the Aristotelian practical rather than Deweyan inquiry for several reasons beyond the obvious one that Dewey is so frequently misunderstood that Schwab may fear he will suffer the same fate if he casts his argument in Dewey's terms. One of Schwab's additional reasons may be that Dewey's dialectical merging of the theoretical and the practical (as is dramatically shown by the grammar of sentences like "for any theory-and-set-of-practices *is* dogmatic which *is* not based upon critical examination of *its* own underlying principles")[17] makes it easy for us to lose the Aristotelian practical in the more familiar, though non-Deweyan, theoretical and thus, as Schwab points out, falsely theoretize problems whose nature and significance are actually practical. Schwab also, I assume, wishes to emphasize some points which Aristotle stresses in regard to the practical—that deliberation be concerned with what can be otherwise, not the already settled; that the activity must lie within man's doing, and the like.

---

[15] J. Dewey, *Logic: The Theory of Inquiry* (New York: Henry Holt & Co., 1938), pp. 105-6.

[16] J. J. Schwab, "The Practical: A Language for Curriculum," *School Review* 78 (1969): 1-24; *College Curriculum and Student Protest* (Chicago: University of Chicago Press, 1969).

[17] J. Dewey, *Experience and Education* (New York: Macmillan Co., 1939), p. 10. (Hyphens and emphasis added.)

But Deweyan inquiry too is concerned with changing the existential situation through considered action. And this action is the solution of genuine problems, the student's problems (in the case of educational inquiry), not the teacher's problems or those thought up by the writer of the lab manual or the textbook.[18] No one ever said such genuine problems were easy to find and handle—problems whose effects on the continuity of the learner and his environment would "live fruitfully and creatively in future experience." The complexities of Dewey's theory of inquiry and of his concept of educative experience have given rise to aberrations on the part of his would-be followers and misunderstandings on the part of his would-be critics. Nonetheless, the Deweyan model stands as a strangely (or perhaps not so strangely) neglected model for educational thought and research.

Whether we deal with Aristotelian deliberation or Deweyan inquiry, the problems are similar despite the differences in approach. In both cases some general principles can be asserted, but these easily degenerate into descriptions of a mechanical routine as exemplified some decades ago in talk about "the steps of the scientific method." Like all other generalizations they are empty until filled with concrete particulars; the procedures of both inquiry and deliberation are to a considerable extent determined by the factors of the particular situation—the precise kind of problem posed, the nature of the subject matter involved, and all the rest.

Since inquiry and deliberation must be made particular to be meaningful, those used in education must be exemplary; they must show the student what it is to engage in inquiry or deliberation of some sort, with the specific instance representative of some genus. In the traditional theoretical fields the organizations and reorganizations given by scientists over the centuries to their amassed subject matters, studies in the "structure" of the disciplines, and similar undertakings have produced a fairly standardized list of topics and even sequences of topics usually covered by educational programs: fungi, algae, molds, mosses; mechanics, magnetism, electricity, heat, and so on. For the practical domain comparable organizational structures are largely lacking. What are the major varieties of practical action, and how are these sorts best sampled in an educational program? Westbury has suggested Aristotle's analysis in his *Rhetoric* as one possible source for such a comprehensive list of topics although he has not yet elaborated it.[19] Certainly without a structure of some kind organizing their subject matter, educators in the field of the practical are liable to fall prey to "current events," procedural gimmicks, or arbitrary fiats.

---

[18] J. Dewey, *Democracy and Education* (New York: Macmillan Co., 1916), p. 182.

The third "nonscientific" characteristic of education with which we began was that it must justify its intervention. Because it presumably produces improvement and not just change, it must show that the state of affairs (i.e., the human being) is somehow better than that which would have eventuated without its interference. Here both the philosopher and the research worker have been much less active than they have in the more "scientific" areas. The educational researcher, if he has abandoned his scientific stance of describing the relations between variables, has usually sought to dodge the normative problem by assuming the role of the technician: "If you choose to procure *x*, *a* is a likely procedure to achieve it; but don't ask me whether or why you should want *x*; I'm a behavioral scientist."

The philosophers, in their turn, have frequently espoused views which appeared to have little utility for the educational research worker or educators in general. If, for example, normative statements are merely expressions of the speaker's emotions, or if they are his attempt to arouse certain emotions in hearers, or if they are simply expressions of personal self-interest, or if they are the imperatives imposed by the victors in power struggles, such statements have seemed of little use in justifying educational operations. "Good" thus became the only four-letter word still tabu in polite circles. As a result, the educator often busied himself with the techniques for achieving unexamined ends or, if he felt he had to consider ends, substituted rhetoric for argument.

---

[19] I. Westbury, "Problems in Search of Attention, or a Discipline in Search of Its Problems: A Discussion of Some Assumptions in a Curricular Theory." A paper presented to the annual meeting of the American Educational Research Association, February 1969, University of Chicago (ERIC: FD 030 180), pp. 16-18; see also R. P. McKeon, "Rhetoric and Poetics in the Philosophy of Aristotle," in *Aristotle's Poetics and English Literature*, Gemini Books, ed. E. Olson (Chicago: University of Chicago Press, 1965), pp. 201-36.

Fortunately, however, recent years have seen two sorts of effort in the normative field which have aroused the educator's interest because they promise to give him some help. One of these has been the attempt to examine the entailments of a serious commitment to some enterprise. Just as a commitment "to do science seriously" demands that the worker record and report his data accurately, publish *all* his findings, and the rest, so we have had several attempts to explore what seriously undertaking "to do morals" or honestly asking oneself the question, "What moral choice should I make?" entails.[20] Although these investigations give the educator some normative ground on which to stand—as Wilson and his group have shown by developing classroom materials for moral education—they suffer from some shortcomings, two of which merit noting here.

First, as their authors are well aware, these inquiries produce "second-order" or procedural directives rather than "first-order" or substantive postulates, imperative or what you will. Accordingly, one may welcome what these studies give and yet feel the lack of specific moral directives. Science can make great progress on the basis of procedural imperatives because the operations are then applied to that relatively firm body of facts called natural phenomena; since no comparable firm body of "moral phenomena" exists (or moral "facts" self-evident as such in the absence of any moral theory), procedural directives are much less productive in ethics. (And, of course, what is true of ethics would be equally true of the other normative fields.) A second charge brought against these second-order or metaethical statements is that they are actually not as value-free or ethically neutral as they allege themselves to be, that their appeal and their utility arise from the fact that some substantive or first-order principles are illicitly smuggled in as procedural.[21]

Acceptance of either or both of these criticisms brings us to efforts attempting to get substantive morals more directly. One

[20] See J. Wilson, N. Williams, and B. Sugarman, *Introduction to Moral Education* (Baltimore, Md.: Penguin Books, 1967); R. S. Peters, *Ethics and Education* (Glenview, Ill.: Scott Foresman & Co., 1967).

[21] E.g., C. Beck, "Educational Value Statements," *Monist* 52 (1968): 70-86; A. Gewirth, "Metaethics and Moral Neutrality," *Ethics* 78 (1968): 218; D. Sidorsky, "Are Rules of Moral Thinking Neutral? *Mind* 77 (1968): 569-71.

popular set of views employs postulation. Here, Hare's *The Language of Morals* can be taken as exemplary if only because the book has been widely used and its pivotal chapter, "Decisions of Principle," has been reprinted in several books of readings for educational courses.[22]

Hare calls for a coherent postulational set of moral imperatives with the derivation of the appropriate actions on the basis of these principles and a review of the probable consequences of so acting. Hare implies that a number of such postulational sets are possible although the postulational process is not so free or unrestricted as it is in the case of mathematics. Since the moral life is lived in the physical world and among human beings and since various characteristics of both the elements are fairly constant, changing very slowly, the postulational sets for morals must take into account these facts about the world and man. Accordingly, certain postulational sets or "ways of life" will be more relevant and applicable than others because they better fit these facts. This general sort of approach, whether Hare's or some other version,[23] does at least offer the worker in education an explicit base for his work even if there are competing sets. Normative issues would at least be clear and could be faced directly rather than ignored or buried under piles of rhetoric.

Over the years since the publication of Hare's book students in education have often asked me why educational philosophers have not followed Hare's suggestion and produced a set or two of this sort, postulates taking account of man and the world as we currently understand them and particularly applicable to the area of education. Educators would then have a choice among several logically coherent and adequate sets of postulates (ethical, political, aesthetic, etc.), the appropriate actions to be taken on the basis of these principles, and the probable consequences of so acting. The educator (research worker, administrator, or teacher) could then adopt that framework most congenial to him or the one best fitted to his situation or purposes. Why, my students have asked, should the worker in educational research, unversed in philosophy, have to build his own set of principles, actions, and consequences in an amateurish fashion rather than take over a set developed by someone professionally competent to see that the set was in fact coherent and the inferences correctly drawn?

---

[22] R. M. Hare, *The Language of Morals* (Oxford: At the Clarendon Press, 1952).

[23] E.g., Henry Margenau, "The Scientific Basis of Value Theory," in *New Knowledge in Human Values*, ed. A. M. Maslow (New York: Harper & Bros., 1959), pp. 38-51; P. Caws, *Science and the Theory of Value* (New York: Random House, 1967).

The reasons are not far to seek. Hare is essentially extending an invitation to build a philosophic system, or at least part of one; and that simply is not done these days. The familiar reasons scarcely need to be rehearsed even in part: the emphasis on linguistic analysis of a few terms, the preoccupation with "science," the view that normative propositions are "nonsense" statements, the attacks on earlier systems charging commissions of the naturalistic fallacy or other errors.

These objections are not at all, however, relevant to the sort of undertaking suggested by Hare. Some of the historical doctrines, if now each is considered merely as a source of related, comprehensive, and mutually related postulates, might well offer at least a good beginning for such an undertaking as Hare suggests. Then time, space, and energy would not be spent—as it was by the original builders of the systems—in "establishing" these principles by attempting to show that they were rooted in the nature of man, in human rationality, or in some cosmic process; rather the effort would be devoted to setting forth in some detail the actions and choices properly following from the adoption of such principles as postulated and to reviewing the probable consequences of those activities. Nor does another standard objection—that such a conceptual structure lies beyond one man's doing—possess its former cogency in these days of team research. One need not cite the shopworn examples of the cooperative effort required to produce the atom bomb or land a man on the moon. One has only to note closer home the centers for study of cognitive learning, curriculum, educational policy, and similar topics where cooperative attacks on complex and massive problems are underway. Most appositely, the work of Wilson, Williams, and Sugarman already cited demonstrates that pooled resources can make possible the progression from abstract philosophic principles to classroom materials.

At a preliminary meeting of some of the authors contributing to this yearbook the question was raised by Michael Scriven and generally welcomed by the rest of the group: What would happen if a few journals of educational research asked all authors to list at the beginning of each article its underlying value postulates, even though the editors would promise in advance that the nature of these postulates would have no bearing on the acceptance or rejection of the article. A request of this sort is probably too revolutionary and too threatening ever to achieve enactment. But it is easy to imagine the greater profundity, thoughtfulness, and coherence that would be attained by both authors and readers and the possibility that bits and pieces might come together in a normative educational structure.

Some kind of normative base must be found if education is to be more than a mindless technology, heeling in the breeze of every whim or rhetorical blast. The rather abortive state of curriculum and counseling—to name only two educational fields—is due in large part to their lack of an adequate normative base; once certain very proximate objectives are stated, discussions and arguments fade out or are converted into oratory. At present, educational psychologists, educational sociologists, and the rest tend to define themselves as behavioral scientists who happen to take their data from that general area of activity called education. But as the preceding pages have sought to show, this stance construes both the "educational" of their title and "scientist" in a very narrow way.

CHAPTER V

Objectivity and Subjectivity in Educational Research

MICHAEL SCRIVEN

Part I

OBJECTIVITY AND SUBJECTIVITY: THE FOUNDATIONS

Introduction

This is a study in the evils of ideology. It is not possible for us to think without some kind of simplifying conceptualization, and in the political, philosophical, and methodological area this is usually an ideology rather than a theory. So this is a study of the evils of oversimplification in ideology, not of all ideologies.

I shall argue that two quite different meanings of objectivity (and subjectivity) have been constantly confused, and the ideology that extols one of these, quite correctly, has been taken to condemn the other, quite irrelevantly. Errors like this are too simple to be explicit. They are inferred confusions in the ideological foundations of research, its interpretations, its applications. It is increasingly clear that the influence of ideology on methodology and of the latter on the training and behavior of researchers and on the identification and disbursement of support is staggeringly powerful. Ideology is to research what Marx suggested the economic factor was to politics and what Freud took sex to be for psychology. Gowin's and other essays in this volume document some of this very well. Perhaps the most pathetic proof is offered by those who condemn *all* ideological discussions as irrelevant to or confusing for good scientific practice and proceed in the next page or

Grateful thanks for a National Science Foundation grant and a Whitehead Fellowship at Harvard for support of some of the work reported herein.

breath—without noticing the inconsistency—to announce their allegiance to some classic ideological blunder.

But general hypotheses about influence are one thing and particular explanations in terms of a specific ideological cause are another. The history of thought is a very complicated phenomenon, and this is not an essay whose principal aim is to advance a hypothesis in that area. Rather, the aim is to identify a number of errors in the currently accepted boundaries between the acceptable and the unacceptable in subareas of educational or, more generally, social scientific research. I am indeed inclined to think that the failure of researchers to see these points has been largely due to the sadly simple equivocation referred to in the preceding paragraph, but whether this is true or not, the consequences of the individual errors have been extremely serious. I attempt to provide more than negative criticism of them; I suggest new directions for research, or at least new emphases in research that I think follow from the cartographic recommendations, and new conceptualizations in terms of which such research can be reported without recreating the confusions discussed here.

The Fundamental Confusion

The terms "objective" and "subjective" are always held to be contrasting, but they are widely used to refer to two quite different contrasts, which I shall refer to as the *quantitative* and the *qualitative* senses. In the first of these contrasts, "subjective" refers to what concerns or occurs to the *individual* subject and his experiences, qualities, and dispositions, while "objective" refers to what a *number* of subjects or judges experience—in short, to phenomena in the public domain. The difference is simply the *number* of people to whom reference is made, hence the term "quantitative." In the second of the two uses, there is a reference to the *quality* of the testimony or the report or the (putative) evidence, and so I call this the "qualitative" sense. Here, "subjective" means unreliable, biased or probably biased, a matter of opinion, and "objective" means reliable, factual, confirmable or confirmed, and so forth. Now it would certainly be delightful if these two senses coincided, so that all reports of personal experience, for example, were less reliable than all reports of events witnessed by a large number of people. But as one thinks of the reliability of reports about felt pain or perceived size, on the one hand, and reports about the achievements of stage magicians and mentalists on the other, one would not find this coincidence impressive.

Yet the urge to simplify is a powerful and legitimate one, and there has been no shortage of those with a philosophical theory to support identifying the two meanings. These philosophical views, however, do not seem to be very satisfactory. The most important of all counterexamples is the case of indirect evidence. If we can test someone's claims on a very large number of occasions and find him extremely reliable, we have good reason to believe him on other occasions when he testifies about events of a kind that we know to exist even though we cannot check them directly ourselves; that is, *unless* we know of some pervasive source of error about such testimony by him that overrides the prima facie case we can build up from his usual reliability. Thus we might well suspect his testimony if it refers to his own infallibility or to his experiences of direct contact with an Almighty God, since there are rather good alternative accounts of such circumstances which he cannot eliminate by introspection and we cannot eliminate by independent test. But if his report simply describes his state of mind and we have no grounds for thinking that he is lying or in some other way misled, especially if the state of mind is a "simple" one such as being in pain or thinking of a colored painting of a sunset, then our previous knowledge of his veracity and sanity gives us good reason to believe him. So the claim is quantitatively subjective, even though it is qualitatively objective, and the two senses cannot be equated. A man's tastes, attitudes and values as well as his thoughts and feelings are quite properly thought of as characteristics of him or of his state at a particular time and are no less objective facts about him than his weight or—to give some simple dispositional properties—his strength or his ability to multiply seven by seven. But of course his thoughts may be private and personal and idiosyncratic and nongeneralizable and all the other terms that mean that they are not objective in the quantitative sense.

Many other complicated philosophical issues came into the disputes of the early days of positivism about the scientific legitimacy of various kinds of statement, including the problem of finding out exactly how one did verify, for example, value claims. But even if there is no way to do so, it does not follow that claims to the effect that a particular person *believes* such a claim are themselves infected with the unverifiability of the claim itself. So the claim that "S believes that S′ values X" is subjective$_1$ but not necessarily subjective$_2$. It is a commonplace of educational research that inventories of the values of individuals are very hard to tie down with any kind of reliability, but the difficulty is not that they are by their nature untestable, just that they are by their nature *often very hard* to test. Yet it would be manifestly absurd, for example, to say that the Christ of the Gospels obviously valued wealth above all else. Since science is not concerned with the search for easy truth, but truth however hard it is to come by, one can hardly conclude from difficulty to scientific irrelevance.[1]

Well, dustbowl empiricism and radical behaviorism had their day, but the task for us now is principally to realize how much damage they have left behind that we have not yet noticed or reconstructed; and that is the concern of the rest of this essay. The problem is not just that the rubble gets in our way. If we cannot straighten out the situation, we are doomed to suffer from the swing of the pendulum in the other direction, a swing which it is easy to see implicit in the turn toward irrationalistic, mystical, and emotional movements thriving in or on the fringes of psychology today. There is much good in them on their own merits, but the ideology that is used to support them is likely to breed the same intolerance and repression that the positivists spread through epistemology and psychology for a quarter century. I shall now take up half a dozen topics and try to make these general charges more specific. The general intent of the paper in redrawing the boundaries of legitimacy is undoubtedly to increase the domain of what has previously been thought of as the (qualitatively) subjective so that many of what have been thought of as illegitimate practices and topics can be tackled. But there are some places where the reverse emphasis will emerge.

---

[1] At this point I shall add that in the early discussions with the other contributors to this volume, I used a different terminology for the two different contrasts or senses of the subjective-objective distinction. For obvious reasons, I had thought to call them the "descriptive" and the "evaluative" senses; but since that very distinction is itself an example of something distorted by the erroneous mythology that I am discussing, I substituted "taxonomic" for "descriptive." In reworking the paper, I decided that I was still doing too much to support an insidious confusion, and so I moved to the present terminology.

Finally, the most important property of ideological papers should be their brevity, and this one lacks even that virtue. Whereas I had originally thought it would be desirable to document each of the alleged confusions extensively from the literature of educational research, my fellow contributors seemed to feel that there was little doubt about the support for the positions I was attacking and that the space would be better spent in making the points of my argument as clearly as possible.

The Background

Without making any pretense at an adequate treatment of the history of these issues, it is worth mentioning some of the elements that seem to figure in that history since it is otherwise difficult to understand either the lacunae in the perceptions of methodologists or the absence of any further need for the drastic countermeasure they were implicitly proposing in the identification of the quantitative and qualitative senses of the objective-subjective distinction.

It is surely in the reaction to introspectionism that we must find the original source of the confusion we are discussing. Here was a research orientation that had been both dominant and deadening; and so it had to be dug out, root and branch, and no matter that some of the branches were fair and flowering.

Moreover, the early days of the opposition movement were extremely successful. So its oversimplifications that might easily have had to be refined in order to yield progress were instead reinforced. And by the time that the leaders of that movement began to run out of successes, they were too old to be able to change. So the next generation, conscious of the deficiencies and no longer impressed by the need for the emphasis, are simply reverting to very nearly the same mistakes that characterized the introspectionists and some of the other nonscientific "psychologists" of the same period. There could hardly be a better case for the utility of teaching the history of research in a field to those who are aiming to do research in that field. Those who do not study history or at least its lessons are doomed to repeat its mistakes. For the social sciences to refuse to treat their own behavior as data from which one can learn is really tragic. There is indeed a danger of coming to treat history as an end in itself, as has happened to a very large extent in philosophy; but if history is written and studied and taught as a record of cases which simply increase our data-base with respect to the task of deciding between several alternative strategies for research, it can hardly fail to improve on a dogmatic and imprecise formulation of the strategies that have seemed successful in the current swing of the pendulum.

Simultaneously with this revolution in psychology, and not wholly independent of it, there was proceeding an equally fundamental revolution in philosophy, specifically in the philosophy of science. And this produced powerful support for the same gross oversimplification. The vehicle was the verification principle, in its various forms and interpretations, running down to operationalism. All forms stressed the notion of intersubjectivity as the criterion for objectivity, not at all in the words in which they were formulated but always in the way they were applied. Just as in the case of psychology, the whole point of this movement was to be seen as a reaction to the obscurantist metaphysics that had got a stranglehold, in this case, on philosophy; a stranglehold so firm that the revolution there was led by scientists and not philosophers.

The same pattern has characterized other areas where the academies became contaminated by concern for doctrine and style and had none for originality; it might be argued that it was not accidental that Pavlov's orientation was psychological and that biology was recently revolutionized by the chemists, linguistics by a philosopher, and sex research by an entomologist.

Just as in the case of psychology, the early days of the new movement in philosophy were successful ones, and for many years there was no serious need to look carefully at the crude version of Occam's razor that had been foisted on the twentieth century.

A third important strand in the movement towards methodological and ontological asceticism reached its full strength later, partly because of support from the two already mentioned; but it achieved power in its own domain and some of this power was fed back to the associated fields. This third strand became known as the "empirical" approach in the social sciences, but it is only significant in some of them; its significance for the others was the support it lent at an ideological level to their related movements. It was of course in political science and sociology that this tendentiously titled movement became dominant and in the other fields that its shadow came to support the weakening hold of crude behaviorism.

It is not irrelevant to look even further—to the remarkable achievements of the physical and biological-medical sciences—for important support for the error we are discussing. For they were *seen* as based on one version or another of this error, although more careful study has shown that they were most certainly not so based. Even relativity theory, for long the favorite example of operationalism, turns out to rest on a deep commitment to non-operationalizable definitions and realities.

There were two awkward exceptions to this harmony of voices in the methodological choir. And their difference led them to be attacked very savagely. One was psychoanalysis and the other was quantum physics. The first really did confuse the issue because the insights of its founder rapidly became the dogmas of his disciples and hence attacking it became necessary but difficult to combine with paying respect to the generally sound methodology of its founder. Quantum physics contributed its own difficulties; the most persuasive interpretation (the Copenhagen interpretation) was ruled out as unacceptable by many of the neopositivist philosophers of science and by the Marxists, who saw themselves as the only legitimate inheritors of the hardheaded hat, roughly because they felt that it was too soft ontologically. On the other hand, the Copenhagen interpretation was immensely successful and could be defended *precisely* by appealing to the need to keep subjectivity out of science, specifically by demanding that we not inject our feelings as to what properties particles *had* to have when there appeared to be no necessity for doing so other than the charms of an earlier physics. Throughout the debate, the language of objectivity and subjectivity recurred.

This has been the background, or so it seems to me; but let us now turn to the issues themselves—to the specific errors that I maintain are connected with the ideological oversimplification to which it is time we gave a name. I shall refer to it as the fallacy of *intersubjectivism.*

The Evils of Intersubjectivism

This erroneous identification of qualitative objectivity with the quantitative notion of intersubjectivity might also be regarded as the fallacy of what has been called methodological behaviorism, i.e., the view that, without committing oneself to any stand about the reality of mental states (an *ontological* question), it is good *methodological* practice to ignore them. (The *radical* behaviorist, e.g., B. F. Skinner, of course denies the *reality* of the mental, as conventionally conceived.) The intersubjectivist does indeed dismiss mental states, but he has a slightly different perspective on the matter. He usually thinks that mental states are only a small part of the domain of subjectivity and that *none* of it has any place in science. So his position would imply methodological behaviorism as a special case. (There are other possible versions which I shall not now discuss. The fallacy of intersubjectivism is not, as far as I can see, related in any interesting way to anything that might be called the fallacy of objectivism—Miss Rand's doctrine.)

The close connection of intersubjectivism with behaviorism is important because in educational research it has probably been in the guise or under the sponsorship of behaviorism that the fallacy has received most support.

It is well to remember, before we get into the forest of abstract methodology, that we are not just discussing an abstract issue with little significance for the practical man. Apart from the reasons given earlier about the very powerful influence of ideology on research directions and hence research payoffs and funding, which affects every citizen, the intersubjectivist fallacy and associated confusions have long since spread from the classroom into the workshop and the public consciousness. References to misconceptions of subjectivity based on this fallacy could be given from reports about wines in gourmet magazines, reports in office supply catalogues, reviews of stereospeaker systems, instructions to news reporters, and texts for door-to-door salesmen. I shall content myself with a recently encountered example from *Popular Electronics* magazine, a hobbyist publication: the context is a report on a new kind of TV antenna.

> The Sensar was tested by comparison with the Pyramidal antenna, and the results given here are subjective. (No field strength measure-

> ments were made.) At a distance of 27 air miles from New York City's Empire State Building, the Sensar provided color signals of the same quality as the multielement Pyramidal on all VHF channels. There was slight ghosting on channels 2 and 4. In the UHF TV band, the Sensar out-performed the Pyramidal and produced notably better color pictures with less grain (p. 82, March 1971).

It is clear that the reference to "subjective" here is to the qualitative sense, since it is incomprehensible except as a mild apology. But then we come to an expansion of what was actually involved in the "subjectivity," and it turns out to be just the absence of a mechanical instrument. As a basis for an apology, *that* only makes sense on the assumption that some existing instrument made of machine parts is always more reliable than a human observer. And indeed, it is very common in the history of the intersubjectivist fallacy that this further error became accepted as dogma in the psychology laboratory. But it is not only an unsupported assumption in general, it is now well known to be false, and one of the areas where we have come to realize its falsity most acutely has been in consumer product testing. For it is here that the gap between instrument measurements and quality has turned out to be the hardest to bridge, and we have discovered that we can quite often get excellent reliability in direct judgments of merit even when they do not correlate with any known instrument reading. In the case in question, it is not accidental that a field-strength meter was not used; they are easy to get and use. The fact is that they do not cover all that is involved in the merit of an antenna's signal; they tell about crude power, but they do not record distortion. And the meters available to measure the two main kinds of distortion run into the problem of calibration in psychophysically meaningful units. We can easily tell you that the Sensar adds 1 percent harmonic distortion and 2 percent intermodulation distortion, but what will that tell you about how the picture looks? The only fact that is important is that those levels of distortion do not produce noticeable deterioration in these circumstances, so you don't need meters. The overall gain in signal strength represents a real advantage only because it *does* show up on the viewer. Further, it is well known that the judgments of picture quality being made by the tester in this report are about as reliable between persons as any practical judgments, which means the correlation is well over .90. So the reporter is quite justified not to bring in someone else to verify them, which he could also have done quite easily. He can rightly say that he knows what they would say just about as well as he knows what they would say if he asked them to check a meter reading. Judgments of picture quality may in fact be slightly more reliable than meter readings because they demand more concentration. In fact, he has given the most objective report that there is any call to give—in the qualitative sense. Yet, because of a hang-up about the other sense, he is apologizing! That shows the power of the intersubjective fallacy, which persuades him that meters would be or might have been better on the inadequate grounds that more people will agree about meter readings than about picture quality. He really knows that's probably not true and certainly not significantly so, or he would use a meter here as he does in other tests, e.g., of the parameters of amplifiers or the insulating properties of a new circuit board material. But the habit of thought is there through the constant pressure of the intersubjectivist fallacy in all methodological discussion. And it is clear from Gowin's article in this volume that the level of thought about similar issues amongst well-regarded writers on educational research is no better.

The cautious reader may feel I am reading too much into the passage quoted. Could one not interpret it simply as saying, in effect: "I did it *this* way (subjective), not the other way (which I *don't* think is necessarily better)." If this were the case, he would be using the distinction between objective and subjective to mark *what*? It is not ex hypothesi a difference of merit (the qualitative sense); it is not a matter of the number of people who would have been involved in the alternative approach—he would have read the meter himself. Are we really to suppose a third *meaning* for "objective," namely, "measured on a quantitative scale by mechanical means"? I suggest it is more plausible to suppose that the phrase in quotes is a secondary indicator of merit via the (carelessly assumed premise) that the interjudge criterion reliability is higher in such circumstances. I say "carelessly assumed" because the assumption makes good sense only if what we're after is decibel gain, which the meter reads. But since what we are after is picture quality, interjudge consistency on meter readings is not going to give us validity.

Disentangling even such a brief example requires some work, but it illustrates the long reach of the issues. Rather than repeat such analyses with every subissue, I shall outline a general critique and make some suggestions about alternative conceptualizations under a series of subheadings referring to applications.

DEFINITIONS

We have already remarked that the chief inheritor of intersubjectivism in the area of definitions was operationalism. The sound element in this approach was the single claim that effective communication requires commonality of meaning and that in the absence of reliable telepathy this can only be ultimately checked by reference to environmental S/molar-R tests. As Mandler and Kessen make clear in *The Language of Psychology*, that does not amount to operationalism at all. Of course, the pressures that led Bridgman himself to introduce paper-and-pencil operations had already undermined the simplest (and probably any distinctive) version of the thesis, as had its self-reference. For how can one give an operational definition of "operation"? If one cannot, the view must either be rejected by application to itself or because it cannot be applied to itself. To sum the matter up, definition should convey meaning clearly, and *sometimes* a useful way to do this is to refer to test conditions ("operations"); sometimes, to take the other extreme, the best way is to define the term by reference to some other terms that we have already worked with long enough to have reasonable confidence in—and these may be wholly abstract. Operationalism, in fact, is either antitheoretical or empty, quite apart from being self-refuting or directly false. And whence came its charm? From the feeling that what could not be seen or done by others either did not exist or at least should not be taken seriously by science: intersubjectivism. Hence concepts themselves—creatures of the mind—were to be excluded as legitimate foundation terms in a definition. (And—a consequence that would have been even more serious for psychology except that the subject was too unsophisticated to notice the lack—the distinction between actions and behavior could not be made.)

But surely concepts could be used in definitions if they had themselves been defined operationally? This sensible recursive approach converts a doctrine about the nature of legitimate definitions into a doctrine about the necessity for at least one operational definition in science, which is not all that the operationalists had in mind. And even this is only interesting insofar as there is some precise difference between an operational definition and a definition with some empirical content, because everyone was already committed to the view that science had to have some empirical *reference* somewhere. And, as we have seen, making the notion of operation exact is not possible for operationalism. It would in any case have been very tricky to see how it was going to be defined so as to make ostensive definition operational, for though teaching someone the meaning of "red" ostensively is an operation, it does not make up an operational definition, i.e., an equation of meaning with the results of performing any *independently specifiable* operation. The operation in question consists in pointing to red objects and saying "red," (and it should include, though this is usually forgotten, pointing to objects that are similar to the first object in all respects except color and saying "not red"). But this procedure will work very well even if all the objects in the universe exchanged colors overnight, since we would still retain our color sense and could easily identify the new red set. The trouble is that this suggests that the ultimate court of appeal is our private experiences and not some feature of the external world. So the only defensible way to define redness, squareness, etc., leads to a definition that makes reference to red things and is hence circular as well as being "subjective" in its reference. Hence even the residual version of operationalism is unsound. This is the record of a desperate attempt, indeed, and only if there lay behind it some very strong feelings could it make sense that good men should have taken this route. The urge to objectivity is legitimate enough; the identification of this direction with externalizing reference in order to achieve multiple witnesses was the unjustifiable oversimplification.

COGNITION

Under this heading I include some of the more basic considerations that are related to epistemology on the philosophical side and cognitive psychology on the empirical research side. Other cognitive considerations will occur under later headings. The intention in beginning with these topics is to show how broadly and deeply the subjectivity issue goes.

Formalism and intuitionism. The long debate about the foundations of mathematics has really centered around problems of subjectivity and objectivity. It is worth considering because mathematics *looks* like the least subjective realm. The formalists felt that it was possible to eliminate all the undesirable elements of subjectivity if they could reduce proof to the application of precise rules to a demonstratedly complete set of tautological axioms; the intuitionists felt that there had to be some further conditions on the rules that would embody their perceptions of the nature of mathematics. It is probably true to say that both sides feel they have been largely vindicated by the outcome of the various lines of research. Without going into it in detail here, it seems hard to deny that intuition comes in in some ways that are ineradicable and significant. The most important example arises from the fact that formal axioms permit nonstandard interpretations and indeed probably must (the Lowenheim-Skolem theorem). Some of these interpretations do not even consist of mathematical entities: hence, since we can recognize this fact, our concept of mathematical entity is not fully formalizable, is partly intuitive.

In the recognition of the validity of a step in a proof, we are normally relying on perceptions of a very simple kind that can be regarded as in the best traditions of intersubjectivism, but there are two special cases that spoil the generality of this view. In the first case, we are facing the first step of a proof that begins from one of the axioms, and, of course, validity must involve acceptance of the tautologicality, or certainty in some other sense, of the axiom. And there we run into trouble; we can't get agreement about which axioms are to be regarded as tautological, or which to allow that are not tautological but are necessary in order to get as near completeness as the Gödel results allow. We cannot forget that Euclid thought his axioms self-evident but there turned out to be options. The Axiom of Choice is a contemporary example. In the second case, there are times when steps in a proof are, or might be, so hard for someone to see that it becomes a matter of dispute whether they are valid. The whole idea of proof is to reduce the steps to such a simple level that "everyone" can recognize their validity. That idea is not quite so easy to carry out in practice, since we get groups of experts disagreeing on whether a step is a "proper" application of the rules or relies on sophisticated perception which can well be regarded as intuition. Cases of self-reference provide one of the most notorious examples, but there are other cases of interest including those of very, very complicated rules, operators, or definitions (like the logical definition of very large numbers) that do not permit any simplification. The latter examples have not turned out to be of any practical importance yet, but may become so with computerization of mathematics.

To put the matter bluntly, it turned out that even formalists have to have intuitions and that some of them turned out to be wrong. So there is no solace in the field of mathematics for the ultimacy of simple external appeals, for intersubjectivism as truth. Of course in some sense this is obvious, since in some sense the agreement amongst many is no guarantee of truth. But we are stressing that it is also not a requirement for truth, that intersubjectivity is neither necessary nor sufficient.

The neopositivists in the philosophy of science, such as Hempel, are fond of giving the concept of proof as an example of a purely logical (as opposed to psychological) concept—a paradigm for analyses of other concepts in methodology. I have discussed it early in this essay because I want to stress the fact that there is no refuge there for the attempt to eliminate the psychologistic judgment of "rightness." Not that the private judgment of one man is to be paramount. The intuitionists in mathematics are constantly debating the best way to express their feelings; they have in fact found it very hard to express their views without falling into exactly the formalists' problems. But the *crude* conception of intersubjectivity, which makes it a logically necessary or sufficient condition, will not stand up. Truth is no more the property of the mob than it is private property; it is an ideal to which we approximate through both introspection and public inspection. Moreover, we adjust each in the light of the other.

Thinking about the mathematical examples also brings out the existence of two interactive properties of intuition—*privacy* and *complexity*. I can often communicate the *results* of my (intuitive) judgments and then *they* become public and open for discussion, but I cannot communicate (fully) my reasons. Otherwise I have reduced the intuition to an argument, and the premises of this may be acceptable without intuition. Thus, either there is a private gestalt perception or concatenation involved in intuition or else it is inseparable from reasoning.

Now, what separates it from mere perception of the kind involved in a meter-reading or symbol recognition and hence necessarily acceptable to the anti-intuitionist, whether formalist or empiricist? It appears to be simply complexity. We well know that the skilled eye can see at a glance what the beginner must gropingly infer, but we do not want "intuition" to be simply a name for all trained perception. Nor can we accept the skeptic's view that it is simply a name for *dubious* perception or pseudoperception. For this falls afoul of the commonality of agreement among mathematicians as to what a nonstandard interpretation is, i.e., as to what mathematics is *not*. Here we have some intersubjectivity about their *conclusion*, but yet there still seems to be subjectivity—because we cannot explicate the reasons for the judgment. Why not? For the same reasons, I believe, that we cannot give a quantified description of each of our friends from which someone who had never seen them could recognize them either reliably or immediately —in short, complexity plus (or, to the point of) nonverbalizability (which is the crucial part of privacy).

Many of our abilities are subject to this limitation. It is no surprise to find that in the kinesthetic domain great athletes cannot tell you what they do in a way that leads your muscles to master the performance to the limits of their ability. But the intersubjectivist is always highly suspicious of such a situation in the cognitive domain. For he believes in "full and open disclosure" as the mark of objectivity, and he takes that to include the reasons for success as well as mere success. Strictly speaking, the grounds for accepting any claim must *ultimately* be intersubjective, at least potentially so, or they would not be grounds. But the intersubjectivist tends to balk earlier, namely, when those who agree on a claim are reacting to a number of features of the situation to which the claim applies, yet we cannot explicate the relationship between these factors that constitutes verification. This is the "complexity" that distinguishes intuition from perception of, e.g., color, and saves the intersubjectivist from inconsistency in accepting the latter and not the former. Is he not still inconsistent in accepting the fact that he can recognize people without being able to say what features control the response and to what degree? His best reply is to say that the criterion performance in the recognition task is independently testable; we check by seeing if the man you say is James Agate carries identification in that name, and so forth. But with the intuition of the mathematicians, we do *not* have independent means of verification; hence we should be haunted by the problems of religious experiencers cross-validating each other and art cliques "proving" the merits of Dadaism by assembling a number of supporters. Our defense—not impregnable—is extrapolation from *previous* performance records by these mathematicians, where their intuitions *were* subsequently confirmed.

We might put the situation this way; if the intersubjectivist simply identifies common agreement about claims with their objectivity, then he runs into well-known counterexamples of pooled error. If he goes beyond this, other counterexamples lie in wait, ready to push him towards the conclusion that "objective" simply means "the best available judgment of the truth" and that "appeals to numbers" are never more than secondary indicators. He has saved himself from the recognition counterexample, but that will not save him from the absolute pitch case where the elite respond to something he cannot check directly, and that is the relevant analogy to the case of mathematical intuition. When experts agree, the agreement cannot be exposed as common error by direct public inspection. Only long-run consequences can be examined by the nonexpert. Doesn't this show that the ultimate basis of objectivity is intersubjectivity? It shows that the ultimate basis of *your* knowing that *someone else* knows something is his doing something that you *can* see (or know about) and that can be best explained by the hypothesis that he knows. But one should not confuse evidence for a claim with the claim itself. Claims may be objective though no one else will ever be able to decide whether they are. Their objectivity is their freedom from bias, their well-groundedness (which does not mean their provable-to-another well-groundedness). Intersubjectivism is thus *ultimately* wrong. But it might still be a reasonable methodological doctrine.

Most examples of its applications to educational research involve grossly simplified versions—for example, the demand that definitions be formulated only in terms of operations. Even the simplest refinement by an intersubjectivist would lead us to modify this claim to the more modest demand that definitions be intersubjectively applicable, and this is not too unreasonable. Of course, we soon see that some technical definitions in the sciences cannot even be understood by just anybody and so we have the restriction of the "subjects" in "intersubjective" to an elite. Now we move to the second refinement which enables the proletariat to distinguish between elites and "crazies" in terms of (publicly) testable payoff. And so on. Intersubjectivism is epistemologically fallacious, but its *sophisticated* forms are often methodologically therapeutic. Our problem, however, is the crude forms.

Subjective probability. Statistics has come into its share of the great debate more recently. It was always felt that the definition of probability was pretty hard to settle but that the dispute had little practical significance. But this has changed with the arguments about Bayesian methods and in particular about the idea that Savage first put forward in systematic form to the statistical community—that subjective probability was crucially connected with the Bayesian approach. A deep analysis of the concept of probability is out of place here, but it would be argued that Savage's interpretation is excessively subjectivistic, just as the interpretation favored by quantum mechanicians is excessively statistical. A better compromise is Popper's idea of probability as a dispositional property of a situation with respect to some precipitating circumstance. And yet that is too physicalistic an account; one must face the old paradox that if probability is a property of things (or situations), then the only probabilities in a deterministic world or subworld are 0 (zero) and 1 since all future outcomes will either occur or not occur. It is this paradox that leads one to be attracted by the idea of identifying probability with degree of confidence or rational degree of confidence, i.e., weight of evidence, in incomplete knowledge situations.

But Toulmin's counterexample is decisive: a doctor who, on the basis of all the evidence he has seen, asserts that a particular patient has a 90 percent chance of surviving an operation can be wrong even if that *is* the correct inference from all the evidence he has seen. For it may be the case that there is other evidence which he has negligently not obtained which makes it quite clear that the chance for *this* patient is nearer to 10 percent, and, if this is so, we would say that he *mis*informed the patient. That can't make sense if probability is merely a logical relation between evidence and conclusion, for having done the calculations correctly makes the probability certain. If we retreat from this to *mere* degree of confidence, with no commitment to rationality, the result is absurd since it denies that the confident idiot at the dice tables is really committing a fallacy in allowing himself to be influenced by the Monte Carlo argument. To say that about the practical significance of probability calculations is to misunderstand them.

Once again we find ourselves in a situation where the "pure" subjective and the "pure" objective solution are equally unsound; we simply have to face the necessity for a solution that is more subjective than we previously thought acceptable, but only because we previously had too narrow a conception of subjectivity. In this particular context, it is quite startling to notice the lengths to which people will go to achieve what they regard as objectivity. Here, it was the frequency theorists who claimed the name; but they never had a plausible answer to the question how they were to make good sense of the idea of the probability of a novel scientific theoretical hypothesis, one of the most typical and important cases of the use of the term "probability"; and never was there any answer to the question how one could call a view objective in the required sense when it made essential reference to the limit of the relative frequency of one type of event in an *infinite* sequence of events. If objectivity here meant intersubjective observability, it was a condition that was not being met; if it allowed calculations and the use of judgment, it was clearly no better off than the despised subjective "degree of rational belief."

In understanding this situation, the historical background helps. Although one could obtain very high degrees of agreement between different judges on the calculation of Laplacean (classical) probabilities, some who could make the mathematical calculations refused to accept them as criteria for rationality (usually because of several known paradoxes). To get rid of this last jot of contamination by the mind, the hope was to move to observable relative frequencies, and the enthusiasm for this move was so great in some quarters that the problems of observing frequencies in an infinite series was simply ignored, or dismissed for inadequate reason. A more straight forward procedure would have been to try to isolate the area of paradox and refine the notions gradually until one could avoid the contradictions, a procedure which many of a less nervous turn of mind did indeed develop. (The same alternatives can be found in the foundations of mathematics as reactions to the discovery of the Russell paradox and its brethren.)

It is always easier to be wise after the event. But these events are not over, and the view I am presenting is not even the generally accepted one. I am not arguing it in convincing detail, but only trying to tie together the various strands in it so that one may find it an overall view worth considering as an alternative to the more usual one. If there is a tone of wonder that others might have chosen other routes in some of this, it is intended to convey the sense of a position from which it is easy to handle the same points as the opposition, without some of the more desperate resources they require. It is not that there is no way to make sense of what they did but only that there now seems to be a better way to make sense of what they saw. The importance for educational research is not just that these theoretical problems about subjective probability, for example, affect decisions about the use of different types of statistics; it is that they add up to an overview that affects other areas as well and affects them profoundly. If we were talking about education and not just educational research, we would have to remark on the importance of these issues in the foundations of mathematics and in statistics for the structure of the math curriculum. And since it will be argued later that research should include much more and much deeper work on the curriculum, the gap is not so great.

Weak and strong knowledge. The paradigm of classical physics has long dominated research in the social sciences including education. There is an interesting way in which this connects with the present issue. We aim for and have often discovered some very powerful generalizations in the physical sciences and we tend to think of these (whether observational or theoretical), along with particular observation statements, as being the paradigms of knowledge. Weaker statements than these are normally thought of as way stations towards these higher forms, to be interpreted as qualified or modest hintings at the same types, as expressions of feelings towards them and hence subjective. The true significance of probability statements cannot be recognized in these terms, however, and there are many other types of statements that deserve better than to be thought of in this way. Let us take a brief look at this area.

If p is a strong knowledge claim, then "I *think* that p is true" can certainly be thought of as a timid move toward p. It can also be rightly thought of as a matter of opinion, a matter of personal judgment. And it is very closely related in sense to "p is probably true," uttered by the same person. So there arises a natural tendency to think that *weaker* claims than p, including probability claims about p, are subjective in some useful sense. If one has adopted a full-fledged subjective theory of probability, then this becomes a *very* natural view. Now consider what we are to say about "the chance of rain tomorrow is 15 percent" or "the chance that this student will make it through this college is 30 percent." These are not modest versions of the full claim (nor are they modest versions of the denial of the full claim). It is implausible to suggest that they are simply autobiographical, subjective in that sense, for they are the professional output of highly trained personnel and if they aren't tied to reality by more than the strands of free association, we should not be paying meteorologists and counseling/admissions staffs. In fact, it is perfectly clear how we go about testing the reliability of such claims; we can do it either by checking the reliability of their producer over a large sample of similar claims, or we can check the plausibility of the inference involved by looking at the data on which it is based. The latter procedure will not work with instinctive judgments of trained personnel, and the former will not work if we have difficulty deciding on what is to qualify as a comparable case or if there are very few of them around. In either case, there is usually no simple rule which will enable us to get from whatever we dig up in the process of checking to a conclusion about the reliability of the prediction (the "external probability," we might call it, since there is an internal probability involved).

This is not a situation that leaves us in any practical difficulty; we all know how to go about checking on such claims if we need to. The crucial point is how to conceptualize the matter. Do we *know* something in these cases, or is it that we simply *suspect* something or have a personal and (qualitatively) subjective *opinion* about something? I think it is implausible to deny knowledge here; yet it is not the kind of knowledge we are used to giving as an example of scientific knowledge. (It is important to see that nothing is being said about probability claims *in general*, because very high and very low probabilities can be well analyzed as slightly hedged standard claims.) The meteorologist who correctly expresses himself thus is evincing some knowledge about the weather that others lack. Since it is knowledge, it can no longer be treated as merely opinion, as merely subjective; these claims are about as objective as many in science, in the qualitative sense. Yet there is some inclination to say they refer to the predictor's inner state and not "just" to the external world. I want to suggest that once again we should be willing to forget the dichotomy of external/internal, subjective/intersubjective, and think of these as claims that *both* require and refer to internal states *and* external ones. If we are prepared to think of such hybrids as respectable, we can not only avoid the necessity of saying that probability claims have a different analysis depending on the size of the probability, but we can find several other hybrids in the neglected realms of knowledge that are just awaiting recognition in order to bloom. Tendency statements, ceteris paribus claims, claims about significant possibilities, analogies, prima facie value claims, approximations, and much good advice are all quite usefully seen as weak knowledge claims. A weak knowledge claim is not a poorly supported knowledge claim, it is sui generis, a knowledge claim of a rather different kind from the usual ones. And from the present point of view, it is likely that most of our verbalizable knowledge can only be expressed in terms of weak knowledge claims and always will be limited to this. This is particularly true in the social sciences and in other fields with a very large number of significant independent variables, especially when memory effects are important.

One would thus expect it to be the norm in history and, if this approach is logically sound, it does something to rescue history from the choice between (usual conceptions of) the Scylla of science and the Charybdis of literature. We can, for instance, answer the question, what do we learn from history, without having to produce absurd or trivial laws, or bare particulars about the past, or murmur mysteriously about deepening our understanding of man. What we most importantly learn from history is a range of possibilities—not of probabilities, not of certainties. And of course these are not *mere* possibilities, of the kind that one has in mind when one says, Oh, *anything* is possible! They are *significant* possibilities, ones that have shaken empires or cabinets before and may do so again for all that we know to the contrary. They are thus most deserving of our respect, and with our knowledge of them we can plan more rationally for the future. If we wish to make a certain outcome *more* likely, then we can try to bring about those conditions which on previous occasions, not demonstrably irrelevant to the present case, did bring about that outcome. If we wish to prevent an outcome, what we can do is to make it *less* likely; we try to eliminate those circumstances which have in the relevant past brought about this result or repeat those circumstances that have previously frustrated it. We rarely have much idea about *how much* effect such actions will have, speaking precisely, but we sometimes know that they have a "good chance," or are a "desperate hope," and to suppose that we shall ever be much better off than that in human engineering a history-conscious world is pipe dreaming. The sad thing is that we could have done so much and done it so much better if we had been willing to learn the lessons from history that are there to be learned, instead of going in search of some Holy Grail whose contents would give us the same kind of predictive reliability in history that we have in astronomy.

The most immediate example is the history of education; it is a gold mine for anyone working in contemporary innovation or dissemination or administration. Not because one can find out from it exactly what will happen if we do X, but because one can find out from it what happened when something very like X occurred before, and hence avoid *some* of the *worst* "goofs." It is full of relevant, though not identical, cases and of good ideas for handling problems like the ones we face today. Dewey the Obscure is revealed more clearly in contemporary accounts of what he did in the teacher training workshops he ran than in his writings; and what went on in those workshops is so embarrassingly similar to recent "innovations" in science inquiry teaching as to make the weaknesses of the national approach to education immediately apparent and pathetic. Its weakness is that we have so completely failed to treat the subject as a serious one that we do not even *count* the losses due to an extremely high level of rediscovery and repetition of mistakes, an attitude which would never be tolerated even in that most backward and traditional field, medicine, let alone in engineering and science. It is true that historians of education have sometimes been their own worst enemies, either claiming recurrences where only resemblances existed, or becoming engrossed in history for its own sake to the exclusion of their social responsibilities. But that is really not excuse enough for the rest of us and certainly not for the system of education as a whole which has not sought, taught, or applied the lessons available. I think this is in no small part due to the lack of excitement many have felt about the possibility of obtaining from history the kind of knowledge we regard as glamorous or even worthwhile. It is against that erroneous paradigm of merit that these words on the model of weak knowledge are directed.

Part II

OBJECTIVITY-SUBJECTIVITY: METHODOLOGICAL APPLICATIONS

Explanation

As one begins to divorce oneself from the hypnotic grip of formalism and its commitment to a split between the mental and the logical, seen as the subjective and the objective, one comes to pay more attention to what was previously seen as trivial, to the mnemonic function of notation and terminology, for example. The latter are no longer seen as mere crutches, but as the vehicles of thought and logic itself. For example, the whole notion of understanding (or insight) turns out, it seems to me, to be precisely a function of the way in which our information is stored in our memories, and hence to be crucially connected with notations and mental images—the "crutches" which the formalist and intersubjectivist wish to discard. In this section we shall explore some of the relatively general features of this approach to questions of the function of theories and theoretical entities—causes and laws and reasons—in explanation of human behaviour and in theories about the nature of explanation itself.

THEORIES VS. OBSERVATIONS

It has never been disputed that abbreviations are legitimate devices for the expansion of a scientific language beyond observation reports and summaries; indeed, it has occasionally been argued that they are the *only* legitimate ones. What is abbreviated is to be combinations of observable predicates glued together with logical connectives, and what they are to do is replace or forestall the need for ontologically dangerous theoretical entities. There are several reasons why this austere approach has met with little success, and it is worth setting them out because the failure of the general program has still left many scientists feeling that they should adhere to it as closely as possible. In the first place, and most importantly, there can be no distinction between observational language and theoretical language as such; hence no way to tell what is a "legitimate" abbreviation. In the second place, there is no reason to think that explicitly defined combinations of properties are always more reliable for communication than less precisely defined theoretical concepts; indeed, complex versions of the former are obviously constructible which will be far less comprehensible and hence less reliably usable than many of the common examples of the latter. In the third place, the aim of theories is explanation, and there are a good many circumstances in which no abbreviational definition can possibly explain whatever it is that we are trying to explain. Fourth, the "logical connectives" turn out to include some very hairy theories of their own, especially the conditional. And, finally, the amount of condensation available by explicit abbreviation is simply not large enough by many orders of magnitude for the economizing needed by the brain. To sum it up, the reasons for theoretical terms are that all terms are theoretical from certain points of view, that there is nothing intrinsically wrong with theoretical terms, and that nothing else can do what they do.

Now, even though this is a fairly extreme line by contrast with the hard-nosed empiricist line that assumed a sharp observation-theory dichotomy and grudged every deviation from the straight "intervening variable" (abbreviations) vocabulary, it does not go nearly as far as the Kuhn-Feyerabend relativism according to which all observation language is *so* theory-contaminated that there is no way in which one can legitimately talk of the disproof of one crucial theory by observations described in the language of another.[2] What is being said here is simply that the notions of theory and observation are entirely *context dependent*: e.g., what are regarded as observations for purposes of constructing a theory of effective school administration (e.g., number of crisis-free days) can be regarded by an individual psychologist as interpretations as far as his theory of group relations or person perception or psychodynamics is concerned. Science does not rest on an ultimate basis of absolute observations; each area in it is supported both from above and below by the network of knowledge to which that area is relevant, part of it more theoretical but supported by independent evidence, part more observational and reliable, as far as the *present* sources of doubt are concerned, than what is being disputed. The strength of a net, unlike that of a building, does not rest on one area's integrity. When we find trouble, in the form of a false prediction, for example, we can adjust the net at many related points in order to minimize the effects. With this model of the structure of knowledge, and with no need to move to the "self-contained pod" model of Kuhn, we can suggest how bizarre a notion an atheoretical approach would be—for each strand in the net is colored theoretical in one light or another.

What *is* a theory? Just a set of claims about the nature of things that serves to explain (or perhaps only to predict) why *certain* things are the way they are. *Those* things, which might themselves be regarded as open to reinterpretation in other circumstances, are what we take to be the givens, the observations, for the present purpose. Of course, when we develop phase microscopy, the frequency analyzer, analysis of variance, surface-effect, computer memories, loudness scales, IQ, the electron microscope, the bubble chamber, etc., many things that were nearly always treated as theoretical become observational; so the process goes both ways. Instead of regarding the observation language as something which must be arbitrarily stipulated or mentalistically defined, therefore, we can retain a useful distinction by using the simple functional notion of context dependence. And thus we can avoid the forced choice imposed by the too simple approach which regards the theoretical as the mental, hence the private, hence the subjective, hence not objective.

---

[2] See also the chapter by Petrie in this volume.

PREDICTIONS AND INFERENCES

There are obviously some awkward elements in the notion of explanation as far as the intersubjectivist is concerned; the whole problem of what understanding is will bother him, since in its usual sense it has to be regarded as pretty highly mentalistic, private, and idiosyncratic. The neatest suggestion in the whole methodological book was the idea of identifying the logic of predictions and explanations. Given that the "logic" of predictions simply consists of inferring them from some relevant correlation plus observed antecedent conditions, this maneuver killed two birds with one stone; it eliminated the mentalistic element in understanding, and it provided a wonderfully powerful test of the adequacy of supposed scientific explanations. The attraction of that double-threat device proved irresistible, and I think this "inferential model" of explanation did more harm to social science than any other methodological doctrine, with the exception of the value-free dogma. It was partly an effect and not just a cause, so one cannot complain that *it* set the subject back by decades; but I find it hard to avoid the conclusion that the views which it symptomatized as well as encouraged cost a quarter century at least. The difficulty with rejecting the view was partly the difficulty of providing a better one, one with fewer drawbacks. The chief opposing view, espoused amongst others by the hardheaded author of operationalism himself, was thought by the formalists to be softheaded; this was the view that explanation is "reduction to the familiar." Naturally, this stank of subjectivity since what is familiar for one is not familiar to another or, indeed, to the same individual at another time. Besides, how was one to define "familiar" in any respectable way? The neatest counterargument came from Popper, who said that scientific explanation chiefly consisted in the reduction of the familiar to the unfamiliar—chairs to atoms, for example.

I think that it is now possible to see how to meet all these difficulties, and I will try to give some idea of the alternative approach. We require a firm grasp of the notion of context-dependent language, some willingness not to corrupt the language of science in order to preserve the particular theory in the logic of science which is under fire, and a decent humility towards the view that scientists are not entirely different from other beings, including children, chimpanzees, and computers.

Let us first salvage some part of the familiarity theory. What is to count as familiar is of course context dependent. But that does not in any sense make it untestable or improper. In order to explain something to his children, every parent knows that he must first find a viewpoint, a model with which the child is familiar, to which he can relate the new phenomenon, or else he will simply not succeed in explaining it. The situation is no different with respect to a class of students at the university as they are receiving instruction in the elements of physics. In short, explaining is not just producing some account which can be memorized; that is "telling about" and it is precisely what is thought to represent *poor* teaching for the excellent reasons that it does not stay in the memory very long and does not transfer to related situations very well. Scientific explanations are no different from any others, except for one simplifying feature which has misled philosophers of science very seriously. There is usually a common cognitive background for the students in a classroom learning science and even for one's peers in the research community. So there is usually no need to pay any attention to the problem of finding a *familiar* model; any model from the preexisting (precovered) literature will do. The intersubject variability is low, and, as we all know, that is just the situation in which we overlook the fact we have made the *assumption* that it is essentially constant.

So we move from "X explains Y to Z" and "Z are science students or scientists" (and Y is a scientific phenomenon) to "X is the scientific explanation of Y" (legitimate) to "presenting X to people quite unlike Z in giving them the scientific explanation of Y" (possibly still plausible, but trouble ahead) to "scientific explanation can be fully analyzed without reference to the audience's comprehension" (false, since X would not be an explanation at all if it hadn't explained to Z, and it would not have explained to Z if they (Z) had not understood it). Couldn't a man discover the explanation of Y and know that he had, even if there was no one else who could understand it? Sure; Z can be the originator of X (a blow for intersubjectivism), but that does not avoid the necessity for referring to his *understanding* Y from X as opposed to discovering a few facts about Y. And that usually means finding some model from Z's previous experience to which Y can be related by means of X. In the rare and undesirable cases where this is impossible, Z has to introduce a new model. He then cannot be said to *understand* its basic assertions—though he knows what they mean and that they are true—but only the other phenomena he can relate to them.

So Popper's remark is just a play on context dependence. The physicist is trying to reduce the familiar-in-one-context chair to the familiar-in-another-context particles: this is just a case of testing a theory on data.

Given that we find some model which we think we can apply to the new or apparently new phenomenon, what do we do next? We show *how* to apply it, and this very often means that we show how we can infer the phenomenon from the model in the light of the particular circumstances or antecedent conditions that obtain. In short, we *reduce* the phenomenon to the model. Now, given the fact that the models to which we appeal in explaining are in fact familiar to the audience to which we address them, this is exactly reduction to the familiar. The so-called inferential (e.g., the "deductive") models of explanation simply emphasized two features of this process, the first being the reduction itself—the suggestion being made, in one version, that this was always a matter of *deduc-*

ing a description of the phenomenon to be explained from the model. In another version (the later Hempel, for example) the relation was taken to include statistical inference to a high probability assertion about the occurrence of the phenomenon, as well as deduction. The other feature on which the focus fell was the model from which the inference was to be made, and this was radically simplified. It was only required that it be a true empirical generalization.

This splendidly simple account was indeed a paragon of formalist analysis and enough to warm the heart of an intersubjectivist. Nothing outside the public domain was involved and, indeed, nothing very fancy inside it.

But it is disastrously oversimplified. There were some extra assumptions being made that were not even recognized by the proponents of the approach and these proved untenable; but, even in terms of what they were meant to stand for, there was no hope. It is an account with very few conditions in it and each of these conditions turns out to be untenable. To put the results generally, this account of explanation includes many nonexplanatory inferences and it excludes many explanations that are unanimously regarded by good scientists as good explanations—and these provide the data with which it has to show good agreement, since it is a metatheory about scientific usage. In addition, its account of prediction was wrong and it incorrectly identified the focus of explanation as a *description* of the phenomenon to be explained and not that phenomenon itself.

Focusing on some points of particular interest in the educational research field, we note that the identification of models with generalizations excludes most of the good ones in the social sciences since they are nearer to being patterns or conceptual schemes than sets of generalizations, but no less capable of generating explanations of particular events than, for example, evolutionary theory, of which the same can be said. Similarly, the restriction to deduction or induction of *at least high-probability* assertions about the occurrence of the phenomenon to be explained wipes out the *overall* matching of a model to the pattern of the phenomena, the procedure which is most commonly used to achieve insight in the sociology, anthropology, and economics, not to mention the politics, of education. What we have to be concerned with is not the predictability of the phenomenon, much as we would like it and should struggle to obtain it, but the superior matching or fit of one account, often involving many subclaims, by comparison with any others and in particular with the null hypothesis-set. The detective knows damn well why this guy killed his wife—because he caught her playing around—but even if he had studied the man for years beforehand, he might well have had to say that this was not to be expected. The point is that the matching shows it to be the best account of any that are possible at all, and as we know perfectly well from many other cases, jealousy *is* a possible cause. Combined with the well-supported view that there had to be some reason (i.e., cause of this type) for a murder, we have perfectly adequate grounds for believing that we have the explanation.

The whole enterprise of finding explanations in history is a matching one, not a quasi-predicting one, which is why the answer to the question, what do we learn from history, is not as simple as we wish it could be.

The crucial logical difference between the enterprise of predicting and that of explaining is that in the former case we do not, and in the latter we do, have the information that the event in question has occurred. Precisely this piece of information enables us to start the business of differential matching between the limited number of known alternative causes of certain kinds of events and the various features of the situation—the modus operandi comparison, in the criminologist's terms. Until we know which outcome occurred, on the other hand, we can only operate off the base-rates, which means that we always pick the nonoccurrence of murder even given considerable evidence of serious jealousy.

One of the most common reactions to the account just given is, we have *not* got an explanation if we can't reproduce the prediction.

Who says so? The coroner and the hanging jury are satisfied enough to act; in the similar cases from medicine and engineering (to stay outside the disputed area), it is exactly the corresponding account that is given and sought as *the* explanation. What this dispute shows is that bad philosophy of science corrupts one's vision of the data; instead of realizing that explanation is a context-dependent notion, that an explanation is aimed at *a particular* gap in understanding, not *all possible* ones, the formalist continues to insist that it is correlative with prediction. For, he says, if you admit that many people might be jealous but *not* kill their wives, then you can't understand why *this* guy did. But our task is not to understand this guy's *pathology*, but just his action.

How can we understand one without the other? The same way that you can understand that kicking this candy-machine made it return my quarter without having any idea *why* it happens or even *whether* this would often work with this or any other machine. You have to get a firm grasp on the idea that the context is what specifies the kind and depth of the explanation that is required. And, incidentally, you need to be clear that having an explanation that will predict the phenomenon still leaves open many other questions that of course *might* be on someone's mind who asks you about the event. There are no complete explanations except by reference to what someone needs explained; there is no account that answers all questions that anyone might sensibly ask about the event.

But could we not introduce a useful notion of "complete," which just requires that an explanation be complete enough to enable the prediction?

But why would that be useful instead of arbitrary and confusing? It is precisely *that* notion which produced the idea that explanations in the social sciences are incomplete, instead of the idea that they were often perfectly complete though they happened not to be able to do other tricks that one would like them to be able to do. You should just say that the social sciences are very good at giving explanations of some kinds, not of others, and not very good at giving some of the types of predictions that we want from them.

Well, but isn't this now just a semantic issue; aren't we just arguing about what to call things? Indeed it is, as must be any attempt to analyze any concept; it will hardly be appropriate to analyze some *other* concept and try to palm this off under the first one's label.

Yes, but I mean does this permissive view of explanation lead to any *constructive* moves? The other one at least sets up an ideal which you agree is desirable and scientific. Indeed it does; it focuses attention on the development of just the kind of theory that will produce explanations of a legitimate informative kind instead of on the kind that must also produce predictions, and since the search for the latter in complex open systems like the social sciences is largely (not entirely) chimerical, this is a constructive and overdue step.[3]

Well, but I'm not sure that there is much point in *having* this weak kind of explanation; maybe the model being peddled is corrupting the better scientific ideals of the social sciences.

OK; that is a point to which it is necessary to speak at somewhat greater length. On it depends any assessment of the merits of theories in educational research. The key to what I shall say can be put simply enough. Explanations are useful when properly conceived because they still have *some* connection with prediction and control and because they are *crucially* connected with other data-handling abilities, notably recall and inventiveness. For explanations are the vehicles of understanding, and these are the benefits of understanding.

It is not too hard to see—I have worked up a fairly good-looking proof of it and called it, rather flatteringly, the *comprehension theorem*—that there has to be a connection between (a) the data storage procedures you will have to adopt if you are concerned with shortage of memory space and/or speed of recall, i.e., practical information storage, and (b) the normally understood concept of comprehension, understanding, or insight.

If you ask yourself what devices and procedures will be adopted in the design of an efficient data storage mechanism faced with short memory space and high speed recovery demands, you will immediately think of the use of radical methods for condensing information. The most obvious of these are generalizations (including, where necessary, rough approximations); they can be used to store an infinity of facts, using one formula plus interpretation rules. At the most basic epistemological level we use codes, concepts, and constancy-cancellation. At the next level of generality above law, we have maps, models, and theories. These can generate many laws as output, hence the description of them as being at a higher level of generality. Or they may simply provide mnemonics which aid in the recall of relevant generalizations (the so-called phenomenological theories in physics). Moreover, the whole idea of using models has built into it the notion of multiple use; for if one system is a model for another, we need only store the details of one plus the encoding rule for the second in order to store the general features of both.

---

[3] Compare this with Krimerman's essay in this volume on causation and prediction of human action.

What will be the result of the use of such procedures, which are quite explicit in science, but also used to an all-pervasive extent in the preconscious levels of operation of the sensor and central nervous systems? The result will be that the system using them will be able to handle requests for information much faster, *including many requests about events that have not yet occurred!* (There will probably be a cost in accuracy.) For the generalizations that apply to the past usually apply to the future (uniformity of nature) and hence allow predictions, or, in the more general case, relations to patterns, e.g., interpolations and analogies. Putting it from another point of view, any system that is concerned with *ongoing* tasks of data storage must try to develop matrices that will not be outdated any more than necessary, because of the savage costs of recoding. Which means developing models that accurately represent the ongoing and interstitial world. (The limit of instrumentalism is realism.)

Now, what is the crucial feature of all tests of understanding as distinguished from tests of "mere" knowledge; what is it that we are supposed to be aiming for in our essay-examination questions (and in the oral exam questions of the research-oriented candidate)? It is precisely the capacity to answer *new but relevant* questions, the capacity that, as we have just seen, we have to build into the brain/computer/student for the most down-to-earth reasons of efficient data storage. So this is the answer to the worry that we may be emphasizing a useless skill with our view of explanation. We are emphasizing the most basic, survival-oriented, data-handling capacity.

But have you not also shown that the emphasis on prediction was, after all, exactly what is involved in understanding?

It is never denied that this is something we would like to have and we have given an extra reason for that, but the crucial point is *where* and *in what way* the predictions come in. The error in the classical account of explanation lies in asserting the necessity for predicting the *event in question* in order to be able to say that you have what it takes to explain *it.* It is precisely *that* requirement that is so neat, and it is precisely that requirement that is false. Of course, in giving any kind of explanation one will be producing assertions, sometimes even general claims, which have predictive consequences; this is often one of the ways of verifying the claims made and—since an explanation consisting of false statements is not a correct explanation—hence the explanation. But it is *not* crucial that these predictions include the event that is being explained. And sometimes "new but relevant" will refer to retrodictions and reconceptualizations which yield no predictions at all, let alone of the event to be explained.

So we have found another way between the path of the intersubjectivist, who wanted us to renounce any reference to the "inner state" of understanding, and that of the humanist, who talks of empathy or *verstehen* but can give us no way to decide between conflicting empathies. We have shown that understanding, properly conceived, is in fact an "objective" state of the mind or brain and can be tested quite objectively; and it is a functional and crucial state of the mind, betokening the presence of skills and states that are necessary for survival in the sea of information. There is nothing wrong with saying, in this case, that we have simply developed an enlightened form of intersubjectivism. But one might also equally well say that we have developed an enlightened form of subjectivism—put flesh on the bones of empathy.

CAUSES AND REASONS

A special case of the preceding concerns one of the most remarkable results of the formalistic influence, the tendency to ignore the reasons that men and women have for their behavior and instead to look for the kind of causal accounts that implicitly deny the rationality of the actors. It seems likely that there were *two* causes of this tendency, if I may say that without self-contradiction, and one of them came from an unexpected source. For in the attack on the reason as a source of action, we find the most formidable alliance in twentieth-century thought—the humanist tradition in one of its more influential and impressive incorporations joining with the whole "scientific" emphasis of the new social sciences. There were other very different and indeed hostile traditions within science and the humanities. But here was one area where the significance or depth analysis cum psychoanalytic approach was not opposing the tough empiricist-formalist-behaviorist group in the social sciences, for in one way or another they were doing the same thing.

There are several strands in the confusions here; there are mistaken analyses of the concept of cause, there are mistakes about the relation between the physical and the mental, and there are mistakes of fact about the evidential support for the various theories that were being applied.

This reasons-causes nexus has many ramifications in social research. The particular paradox of Marxist thought here (especially with respect to social value assertions which they wished to explain away, but also to make) has been quite well analyzed by anti-Marxists; but new attacks on the "capitalist" social scientists have only just begun to appear. The revolt by the radical psychologists at the 1969 convention of the American Psychological Association is a good example. This took the form of disrupting the meeting devoted to "explaining" unrest on campus and protesting that it was morally and scientifically objectionable to put forward the kind of factor analysis—in terms of superficial institutional variables like size and location—that is usually offered as explanatory when the revolutionaries are offering perfectly clear and quite different reasons for protest. It is clear that these two kinds of explanation may, speaking generally, be related in various ways, ranging from straight incompatibility through mutual support to complete independence, and it's quite clear that no one on either side at that meeting had seriously thought out the relationship: there is essentially no mention of it in the transcript. Each side had favorable feelings towards one of the approaches and was skeptical about the significance of the other, and each side, it seems to me, really *was* being underrated by the other. But there really is no general formula for resolving these disputes; one can only get down to the hard work in each case. Causes are simply single-factor explanations, often the precipitating factor of interest in a complex situation—and a reason is often exactly that. Attempts, such as H. L. A. Hart's, to give essential differences have all failed or degenerated into arbitrariness. Whatever the predisposing or enabling factors on a campus, which may indeed emerge as significant correlations in the matrix, their existence in no way disproves the assertion that the reason for the trouble was the war or racism. At most they may be an interesting supplement. On the other hand, the mere feeling of anger towards the war in a number of those who led the protests does not even show that this was the real reason for their own action—they may have been deceiving themselves—let alone the reason for the actions of the group as a whole. One simply cannot show this by statistics or interview data of anything like the kinds that are available.

There were associated grounds for complaint, e.g., that this kind of investigation tends to give the impression of a put-down ("I'm going to find out the *real* causes of your behavior, which you *say* is a protest against injustice"), and that it involves gathering data which could incriminate students without taking adequate security precautions with the records. These raise somewhat different issues, however, more relevant to the subject of the next section.

The point just discussed illustrates only one aspect of the lack of concern by psychologists, especially educational psychologists, with the function of reasons in explaining behavior. We might say it illustrates the tendency to assume that reasons for behavior usually are or mask rationalizations. Or we might simply regard it as yet one more example of the urge to dismiss what is not publicly observable—intersubjectivism.

An equally serious failure has concerned research on the actual strategies of reasoning and ways to improve them. We are beginning to see some work on this now, but the psychologics proposed in many of the science curriculum projects as an analysis of the process of systematic scientific thought are staggeringly a prioristic and naive. The same can clearly be said about the higher language-handling skills referred to in "propaganda analysis" and "critical thinking" and the "new social studies" and some English curricula. We shall look back on these as inexcusably primitive for a period when science was fairly sophisticated and scientists at research were readily available for study. This certainly appears to be another cost of the fallacy we are considering, for surely one would expect educational research to *focus* on good reasoning if no antirational ideology intervened.

Finally, however, one must come back to the logical questions. What analysis of cause itself (or of reason or of goal) is to be found in the texts that guide the educational researcher? Not *one* that is immune to common counterexamples from the educational field! The most popular suggestion adopts a modified sufficient condition analysis (meaning one with some kind of background condition or ceteris paribus clause). But no such analysis can handle cases of overdetermination where more than one sufficient condition is present but only one operates in fact for it would identify both as the cause. And it is easy to show that on any modified sufficient *or* modified necessary condition analysis, most effects are, *per impossible*, the cause of their own causes (roughly, since the latter can be inferred from the former-plus-background conditions).

Unfortunately, it does not appear that any satisfactory reduction of cause to acausal notions is possible, contra Hume. The excessive anxiety to perform such a reduction was one more symptom of the urgent antimentalism or intersubjectivism of early twentieth-century Anglo-American thought in epistemology. The first question that should have been asked is, Are we having trouble in actual scientific research agreeing as to when given data establish a causal connection? If not (and I suggest the answer would have had to be negative), then there is prima facie evidence of the only kind of intersubjectivity we are entitled to demand—that of common understanding. Our task must, therefore, be to see if we can find illuminating ways to convey that meaning which scientists are apparently sharing, *whether or not* it turns out to refer to some mentalistic or indeed anthropomorphic notion. It was mere meddling to insist that the way to objectivity could only lead through an analysis referring to bare correlations between observables. The intuitions of scientists, as of mathematicians, can provide a good workable basis for a discipline even if they are *necessarily unformalizable.*

Evaluation

It is hard to tell when the last breath has actually been drawn by horses that should have been dead long ago and have certainly been beaten enough since then to kill them twice over. The special problem with the myth of value-free social science is that attacking it has now begun to be a matter of hopping onto the bandwagon. All sorts of reasons for killing off the myth are being offered that are well calculated to lead to a reaction that will ensure its perpetuation. The fact that scientists have to decide which problems to investigate does not show that the scientific investigation they conduct is one whit contaminated by value judgments that originate within science and hence does not show science cannot exclude value judgments. It is, however, certainly relevant to the quite separate (and never disputed) claim that science has social effects and that particular scientists have to concern themselves about their social responsibilities.

I think the best way I can sum up the view I would like to recommend here is this: there are several very different types of value claims (or uses of the term *value*); all of them exhibit quite clearly some properties that incline one to call them subjective, but all can be shown quite conclusively to be essentially, i.e., in principle, entirely—indeed, even *wholesomely*—objective. Personal-value claims (for example, "I value . . .") are indeed quantitatively subjective, but, on the other hand, they are obviously simply dispositional claims about an individual and hence open to objective investigation in the same way as are claims about the edibility of yesterday's lettuce.

The big fight is over what we might call real-value claims ("freedom of speech is a more important value than sparing a few feelings") and particularly the subspecies of *moral* real-value claims. A frontal attack on the skeptical position seems well justified here and consists in arguing that (a) any trained scientist in any branch of science, no matter how pure, has to make and defend real-value claims (subspecies; methodological) when acting as a scientist (e.g., "this is a better theory/experimental design than that"); (b) many scientists in applied fields, of which education and psychotherapy are prime examples, must, *qua scientists*, make moral decisions and incorporate moral decisions in their concepts and theories (e.g., in defining the measures of improvement in treatment/education); (c) the subject matter of ethics coincides with part of that of social science and hence can perfectly well be regarded as a branch of the latter.

Only the last point deserves amplification here. The introduction long since of the concept of implicit function into the repertoire of anthropology, the addition of role theory, game theory, welfare economics, some jurisprudence, political theory, and survey methodology *more* than covers the field of ethics. The resulting secular conclusions about rules and attitudes for interpersonal and intersocietal relationships have to be regarded as a basis for rational action until someone can show they are superseded by a higher authority, and of course that doesn't mean *any* level of "higher authority." It is a remarkable achievement of a muddleheaded nation to have persisted in the belief that there are pretty good, i.e., presumably objective, reasons for democracy (or freedom of speech, etc.) but that ethics is essentially subjective. Democracy would normally be regarded as being a system of government whose institutions are intended to achieve the equality of rights of its citizens. And that is not only *a* moral claim but the *only* moral axiom necessary to develop a full-scale moral system covering every point that the traditional systems cover. If democracy is defined in any nonmoral way, e.g., by reference to government by majority, universal adult franchise, etc., then the process of giving arguments for it must immediately involve a moral premise. If this is taken to be arbitrary, then democracy cannot be taken as objectively defensible; if nonarbitrary, then at least one ethical principle is objectively true. To take seriously the idea that no form of government is objectively better than any other, for a given society, is a sign of acute philosophical cramps. To think that ethics must assume the validity of stronger claims than this is an error based on failure to attempt it without stronger assumptions.

The above is a pragmatic view of value, and the most common objection to such views is raised by asking where the ultimate ends or "goods" come from. Perhaps the most satisfactory answer to this is to offer these alternatives: *either* there is no need to assume any ultimate values any more than we have to assume there was a first day in the history of the universe; *or* the ultimate (but not highest) values are the biologically immutable needs of individual reasoning beings. The latter need not be thought of as necessarily valuable in themselves, only as desired, i.e., valued, in the first instance. That is, they provide an epistemological but not incorrigible starting point for the eventually self-corective value system.

The values field is a notable example of the effects of an oversimple identification of objectivity with whatever achieves interpersonal agreement. How many times does the beginning college student, long since corrupted by the cognitive culture, listen dumbfounded to serious talk of a moral position being "objectively right" and then propose the absurdity of this on the grounds of the fact that there are extensive and long-unresolved disagreements about moral issues. No other explanation of *moral* disagreements has ever occurred to him. We ask him how he would explain disagreements between supporters of different football teams, theoretical physicists, or political candidates, and he produces several possibilities. Why the difference? I think—on the basis of hundreds of hours of discussion—partly because of a failure to distinguish the subjectivity of personal-value statements (quantitative) from the subjectivity of real-value claims, e.g., moral assertions (qualitative). And partly, too, from a belief that ethics is the preserve of religion—a belief supported by the powerful alliance of the churches and the empirical school of social science which found its strength in isolation from values (a situation richly deserving an explanation in terms of causes, not reasons). Yet he will not dismiss democracy—or penalties for murder or drunken driving—as a mere local option.

It is fascinating to hear a Russian trying to understand this divorce of values and social science. I think we would share his amazement if our senses had not been dulled by proximity for so long. It is not enough to realize, as we are beginning to do, that the social sciences had an enormous effect on moral issues just by deciding to cut them out of the fabric of legitimate scientific concern. We must also see that the reasons for that excision were simply fallacious and may even have been acceptable only because of unconscious identification with the value system of the society which was thereby protected. That suggestion is insulting, and it could easily be repudiated by showing that the arguments for value-free social science are so good that they were probably accepted because of their apparent merit. But my present view is that no such case can be made, and that the return to the historical sources shows that Weber, for example, was affected by significant considerations of the political climate of his time that have been essentially irrelevant to the scene for many decades.

Educational research on values has been marred and constrained by the intersubjectivist fallacy perhaps more—and with more serious effects—than research in any other social science area. Other essays in this volume illustrate the changing tendencies today.[4] See under major heading, "Education," below.

Investigation

THE LOGIC OF DISCOVERY

The effects of intersubjectivism on the concept of research itself have been very serious. The most serious of all has been the passionate support for the so-called hypothetico-deductive model of scientific inference—roughly the view that research consists of genius plus statistics. You need to have genius to get ideas that are worth considering—and there is nothing to be said of a logical or pedagogical nature about that—and the process of scientific inference comes in only at the next stage, where you have a hypothesis to test. You deduce consequences from that, compare them with the results of experiments, and use some decision rule to decide whether to accept or reject.

Maybe this was a useful reaction to some other view that was current years ago, but on this one my patience gives out early. The crucial consequence of this position has always been that there is no "logic of scientific discovery," because such a thing would consist of rules, the application of which would automatically lead to the generation of new and interesting hypotheses. Since everyone knows that great discoveries are not made by rule, this possibility is ruled out, and it only remains to examine how logic *is* used, namely, in the testing procedures. This view of the matter, chiefly due to Popper and his school, has received a great deal of support from eminent scientists which either shows that they know a good account of their own activities when they see one or that they rather like the genius part. Anyway, it is a pernicious doctrine on both logical and social-educational grounds. Let's first look at the alleged contrast between logic and something else. The classical kind of logic, the type that can be highly formalized and in which the rules of inference can be explicitly and precisely stated and routinely applied, is mathematical logic. The suggestion is that there is something crucially different between this kind of enterprise and science of the inductive or empirical kind.

---

[4] Note especially chapters by Thomas, Ecker and Kaelin, and Krimerman.

In the first place, it won't do to argue that all mathematical/logical statements are tautological and for *this* reason different from scientific ones. Not only is the antecedent false (axiom of choice, etc.) and also the consequent (mathematical physics), but the point isn't interesting since we are concerning ourselves with the *process* of getting to these propositions, not the nature of the *product*. Second, it is clear that the existence of rules of inference in mathematical logic does not mean that the process of discovering important results has been reduced to a rote procedure there. Applying the rules *may* indeed lead you to some exciting result, but this is fantastically unlikely and so not fruitful as a procedure. Moreover, it could be done in science. So the contrast cannot be maintained between the logic of deductive and inductive systems; it can only be maintained between the logic of discovery in both, which is said not to exist, and the logic of justification, which is said to exist. Popperians often deny the existence of a logic of induction, and the reason just given suggests they would often have to conclude that there is no logic of deduction either.

But is there perhaps another difference between the empirical and the analytic domains which we are overlooking? Well, some very simple logical domains can be *completely* formalized, but this is provably false in general and provably true with some simple empirical domains, so that distinction can be ignored. What can the logic of deduction be, to which we can contrast the (allegedly nonexistent) logic of induction? If there is no contrast, it can't be the case that only one exists.

Well, what *is* this logic of justification? It is certainly true that there is a great deal of *use* of what we would normally call *logic* in the process of justification. The concept of proof is usually said to be an effective one, in the technical sense that one can decide by a finite procedure whether any given proposal is a proof. But once again the inscrutable element of genius has to come in to give us the idea for the proof, since there is in general no routine procedure for constructing one in either the physical or the mathematical domain. So far, then, there is no difference between the fields. One may just enter a caveat to the effect that proof is not really an effective notion, though proof within a system often is; but the point is not essential here and usually upsets formalists acutely. A far more serious problem does arise here, though. There are some cases where one can give a strict proof of an empirical proposition, e.g., by deduction from others, or comparison with data if it has finite scope. But almost always the "logic of justification" can only proceed by showing that one account of the data is better than *any others we can think of*. Previously well-justified beliefs are amongst the first to fall when scientific revolutions occur. So the logic of justification itself *depends completely on* the allegedly nonlogical inspiration that yields the competitive accounts (for example, that the sample was biased economically, because . . .). So the only difference is that we apply the logic of justification *after* the bright ideas are in, the logic of discovery *before*? But there *is* no logic of discovery!

Indeed, there *is*. The whole process of proving that there is not is an elaborate persuasive definition of the concept of logic, plausible, indeed, but leading to results that only reflect a subtle error in the semantics. When we describe a person or an argument as logical, I suggest we are not using the term in the way that the Popperians eventually pin down; we are not at all restricting ourselves to the very limited range of effective procedures in a highly formalized language. We are talking about systematic rational approaches to problems and arguments, or to the performance of those who have developed skills that correlate with the results of such approaches. And that can certainly be said to apply to the procedures that a mathematician uses in searching, *either* for a proof *or* for a hypothesis which he may subsequently try to prove. Equally well, it may be said to apply to the thought processes of the detective, the historian, and the physicist, always granted that no such approaches *guarantee* a successful outcome and that other processes, such as free association of an inexplicable kind, *may* produce the answer. That is, there is a logic of discovery even though it is not a set of rules whose routine application *always* does and *alone* can lead to the required discovery. What those rules are varies from area to area, unless specified very generally. At a middle level of generality, in searching for the cause of something (X) in a highly deterministic domain, we might describe the logical procedure as:

1. List all known possible causes of X; call them C_1, etc.
2. Check on the actual presence of all these C's.
3. Let those present be the set C_1, C_2.
4. List the known modus operandi of C_1 and C_2: call them c_1, c_2.
5. Check on the presence of these modes.
6. If only one present, c_1, then C_1 is the cause.
7. If both present, check on termination time of the causal chains.
8. If times different, the first to terminate was the cause.
9. If times same, event is jointly caused.

Step 8, for example, is not analytic, just a good bet. Following this set might reasonably be called following the rules for a logic of discovery. It does not eliminate making good observations and having good premises, but no logic eliminates these or making good calculations which seem to be essentially similar.

The reason for arguing for what might seem to be a mere rechristening of the situation as seen by the anti-inductivists is that it markedly affects all approaches to research on creativity, on teaching scientific method, and on any specific scientific problem, since the two approaches have a quite different view of the importance of what can be done systematically. If one wants a Nobel prize, there might be some case for thinking that the logic of discovery won't get one there; but even that is implausible in the light of the details of the double-helix story. Yet breakthrough originality is not the typical requirement in scientific research and nearly all of it can be handled very well by what there is of a logic of discovery. Certainly, we need a great deal more work on it; we need systematic attempts to inventory the models that one should have clearly set up as possible templates, ranging from mere checklists of possible causes to the set including field models, mechanical models, and quantum models. I think that the desire to avoid getting into the mental life of the scientist once again led to the rejection of good and sensible approaches on doctrinaire grounds. Work on the heuristics of research has had one or two supporters, but it has really been a backwater, whereas it should be the source of much of our curriculum material inspirations.

We need to make our picture of the work of the scientist more realistic even if it means making it more like that of the auto mechanic. Precisely that similarity may be what has led us to hide the simple model of research for which I have been arguing here—eliminative inference to the best explanation. At least this can make sense of "teaching the discovery approach," otherwise a contradiction. It is not too surprising that what Sherlock Holmes used to call a deduction turns out to be quite nicely reconstructible as exactly that, once the spell of thinking of induction as something quite apart is broken. Induction and deduction differ only in whether an external probability is introduced in the course of the inference and the trained scientist is quite good at estimating just what that probability is; that is part of what we have to learn how to teach in teaching the logic of induction and discovery. Like much of values education, it has been underground education.

THE NATURE OF SCIENCE

Much is implicit in the last two sections about how science itself should be conceived, in particular the emphasis on the continuity of the processes of valuation and inference inside it and outside it. It is perhaps worth adding a little more. The continuity of the sciences with the other subjects that attract our interest is really important, and the continuing use of the old curriculum divisions does not help to get this across, with its effects on vocational choice, self-image, and the way in which investigation is conducted. Out of fifty possible examples, take something like the hobby of shortwave listening. This is a pretty good example of a science—in every respect that matters except the prestigious name—and it's a whole lot more interesting to some students than most of what comes in the texts on the standard science subjects. Reading the club bulletin or a standard report of the kind that the hobbyist sends in to a station engineer to inform him about distant reception is about as hard as reading the Greek alphabet and illustrates many of the points about terminology, observation skills, theorizing, and so on, that I have mentioned earlier. It is useful to try to sort out the different social and logical elements in the decisions to call the study of government *political science*, that of household management *domestic science*, and a religion *Christian Science*, while leaving history, jurisprudence, and architecture in the nether regions. It is, for example, very hard to avoid the conclusion that some areas are left out because of the lack of verbalizable generalizations—good teaching would be an example. Polanyi and others have argued that a large part of science itself (and a crucial part) is not verbalizable. Are we, then, discriminating on some inappropriate grounds? I leave the question open; it serves as a good example of the way in which jurisdictional issues—semantic issues, even—can serve to relate the area of research to the area of teaching.

Education

We have covered many topics that affect the conduct of education, but in conclusion it may be useful to add a note or two and to say something more about possible directions for educational research at a more specific level. But I stress that I think the preceding, if it seems persuasive, should have its major value in loosening up some of the rigidity in our conception of what we are about and should be about, in directions which I do not pretend to foresee. What I do see is ideological reasons being given, reasons that are in some cases definitely the decisive ones in determining research action, support, and interpretation, for not doing some things and doing other things, when it is clear that the reasons are completely fallacious or greatly oversimplified.

Once the force of the intersubjectivist pressure has been somewhat reduced, it becomes much easier to avoid the extremes that are often adopted now to preserve some kind of independence for the subjects that are neither simply sciences nor simply humanities—history being the most serious example. That we have *two* separated cultures, in Snow's sense, reflects an ideology of science, indeed of knowledge, that we must take great care not to encourage. It is unfortunately true that a more serious problem than Snow recognized also needs treatment by the educational system, and that is the failure of the scientists to learn enough about scientific method to act like scientists when faced with materials outside their own specialty, and the corresponding failure of the humanities faculty to show much sensitivity to considerations of humanity when they arise, for example, in concern for the rights of students to blind, cross-checked grading. But then this has been receiving more attention lately and does not need to be labored. More basic is the need to discuss and defend whatever evaluative procedures are used by the faculty.

Some of the specific areas that come into focus as needing attention, with the reduction in the strength of the taboo against values as a legitimate subject for *normative* study, include realistic presentations of the arguments for the major alternative value positions on social as well as international issues, and some real attempts to use anthropological and historical data to increase flexibility in attitudes towards other systems of values, not at the "how quaint" level. As soon as one lists this kind of an entirely obvious extension of the current emasculated curriculum, it becomes clear that a very serious dimension of educational reform is still not receiving enough attention to support such an extension, namely, parental education. Without it, there will be the same troubles here that there were with sex education, and with it there is an added payoff in the education of the parents.

The topic of objectivity and subjectivity nowhere comes closer to home in the curriculum than in the development of a self-concept. For that is so closely affected by others' concept, of us as to merge both senses of objective and subjective. This area of education has been largely left to a subgroup of the affective education people, to third-world studies, and to some enlightened counseling/guidance people. It is simply a prime area for more research and practice, but one where we cannot bring to the work a bias against private experience.

Turning to the vocational side, the implications of the questions being raised above about the difficulty of distinguishing science from other skill/knowledge areas are certain to suggest that the usual glamorous nonsense which is used to support the difference in the wage scales between technicians and scientists needs to be blown away. In the same way, the aims of educational research need some proletarianizing, and some serious concern with teaching manual skills in a more effective and respectful way would be one of the changes that one would like to see. Even further down the status hierarchy, we hope to see educational psychologists working on improving shorthand and typewriter keyboards. Why are not those and speed reading, and memory skills, and calculating short-cuts, and "speed-speak", and "fast-think" the objects and much of the work in the learning area providing the chance for practical payoff as well as theoretical gain? In every dimension one is faced with the list of things that have long needed doing and that have been kept out by status considerations of the crudest and least rational kind. Why do the journals not allow for the publication of ideas as well as articles—paragraphs that might suggest more than would the most majestic unread monograph? Why the fetish for documentation? If you find something interesting in this article and want further reading or remarks on it, write to me and I will send you the most up-to-date reference I know about at the time I reply. I judge from past experience that this will not be an overwhelming burden, and it's a far more useful service than listing fifty dated references or five for every footnote, as these topics demand. Where are the professional journals that are concerned with providing a forum for and obtaining student comment on articles about teaching practice and evaluation, or getting anonymous comment about the operation of the USOE and other agencies, or, in general, better evaluative feedback aimed at improving education? The countenance of educational evaluation has radically changed lately, as we come to see that it is partly social and moral criticism and not just tests-and-measurements exercises. It still has not been sufficiently used on education itself and its research arm in particular.

Conant once defined the sciences as the cumulative subjects. No one who has tried to live with a set of new and expensive school buildings or curriculum materials will believe that education is a serious science—and of course the usual response to the first example is to say that it's not the business of educational research to build schools. The response is irrelevant. It most certainly is the business of educational research to evaluate them or to prove they have no effect on education, to devise ways to get those evaluations into the planning of new schools in the future, and to think of new ways to specify the designs for them. That is one example of what we need to do to show we are trying to develop a cumulative subject, not a parade of fashions. I hope there are twenty others in this article. I am sure there will be another twenty in the minds of the first few readers. I hope they will tell me about them; if they do, I will send back a list of the comments I have had from other readers, destructive as well as constructive, and publish them later if I can find a way to do it. There are many better ways to use our public utterances and publications, and we should be exploring them. We can't afford the luxury of multiplying mistakes through ignorance of others' work or through failure to explore alternatives. One small way to move forward is to try something different when we can. Like this. Is it not a long way—a long way *down*—from discussing objectivity and subjectivity in educational research? I think not; I have argued that our hang-ups on that issue have affected the whole shape of education. If rethinking them didn't lead me to suggest some new things, I wouldn't believe that and neither should you. Suggestions that are manageable, potentially useful, easily evaluated, novel, and cheap might be said to be a long way *up* from much educational research of the past. That this is so reflects profound ideological confusion of the kind this paper has discussed.

CHAPTER VI

Behavioral Definition of Educational Objectives: A Critique

IRA S. STEINBERG

Introduction

The assumptions in the behavioral definition of educational objectives are rather straightforward. The student is to be educated or schooled. Those responsible for schooling him are responsible for setting the objectives of his schooling. To avoid vagueness and ambiguity in the setting of these objectives, they should be specified in terms of the intended testable outcomes of instruction or teaching. Ideally, a proper behaviorally defined objective (or BDO* for short) should anticipate the question, "How shall I know if the student has met the course objective?" by specifying the course objectives as the answers to such questions. Objectives are to be stated as intended performances of the student after instruction. Thus, for a simple example, the fuzzy objective "to have the student understand algebra" would be replaced by a set of objectives of the form "the student solves equations of the form. . . ."[1]

It should be easy to see why this way of defining educational objectives should appeal to the empirically oriented researcher on teaching, classroom learning, and/or curriculum development. The empirical researcher must specify instructional objectives or elements of curriculum in forms amenable to observation and testing no matter how ephemeral they may be as given to him.[2] If, for example, he is to say at what age a student is likely to benefit from instruction in a given subject, he must be able to specify the student performance which will serve as evidence of the success of instruction in that subject. If he is to say what techniques of instruction are most effective in improving student learning in a given subject, then he must be able to specify the sorts of performance which will serve as evidence of learning in the subject.

---

[1] This sort of specification of behavioral objective is perhaps best exemplified in Robert F. Mager's *Preparing Instructional Objectives* (Palo Alto: Fearon Publishers, 1962), pp. 14-18.

* BDO is used to denote a behaviorally defined objective (singular) or behaviorally defined objectives (plural).

There is a sort of symbiotic relationship between the schoolman and the educational researcher in the business of BDO.[3] The researcher may be able to avoid the charge of arbitrariness in defining the objectives of schooling in operational terms, if the schoolmen are, in effect, doing it for him or are helping him to do it. The schoolman may borrow an aura of scientific rigor and professionalism for his evaluation of programs and personnel if he can claim to have adopted BDO as specified in "scientific research" on education, especially in research on teaching and curriculum development. Indeed we might even get the idealized close cooperation and communication between practitioners and researchers concerned for the improvement of learning and teaching. Whether such communication is a good thing depends, of course, upon the value of what is communicated.

Good Intentions and Accountability for Effectiveness

The main practical concern in BDO is for the improvement of teaching and learning in objectively recognizable and measurable terms. If course objectives are stated clearly as what students should do as a result of their participation in a course, then their performance provides evidence of, or a measure of, the success of the course. To the extent that courses taught by different teachers to different students may share the same objectives, the clear specification of these objectives in terms of student performance permits the comparative study of student success in learning and of teacher success in teaching. Such clear specification of what is to be learned and of what is to count as evidence of learning appears necessary if one is to control for the variables in teacher characteristics or teaching characteristics and learner or learning characteristics in order ultimately to speak with any confidence of the optimal matching of these characteristics for the improvement of teaching and learning.

---

[2] It is not surprising that the major impetus to the Bloom taxonomy studies comes from those primarily concerned with educational testing. See Benjamin S. Bloom et al., *Taxonomy of Educational Objectives: The Classification of Educational Goals: Handbook I: Cognitive Domain* (New York: David McKay Co., 1956), pp. 4-5.

[3] See, for example, Will French and Associates, *Behavioral Goals of General Education in High School* (New York: Russell Sage Foundation, 1957).

If people can be gotten to agree on what is to count as evidence of learning in a specific sort of course, the separate specification of what is to be learned would appear to be superfluous and redundant. For comparisons can only be made by reference to evidence. However many and disparate the objectives we might claim for our own versions of a given course, if we are prepared to accept the same tests of student performance as adequate measurement of the success of the course, then our differences would seem more rhetorical than real. In admitting the adequacy of the tests we should be saying, in effect, that for all we can tell there is no difference in our stated objectives; our objectives are given meaning by the tests as we affirm their adequacy. We say, "We are trying to get our students to do the sorts of things required of them in the tests."

Of course, it is conceivable that people can agree on a test which measures proficiency, say, in a range of areas and where it is considered legitimate for a course to focus on preparation for performance in only some of the tasks that such a test might require of them. Courses, then, may differ in content and objectives, but objectives are still defined in terms of performance on the measurement instrument. We might say that the courses are virtually, though not actually, the same (i.e., selections from the range of potential BDO are deemed legitimate and thus *virtually* equivalent representations of the course). On the other hand, we might note (if it is a fact to be noted) that different people have chosen different BDO, that the parts of the test that embody the various BDO appeal to their several partisans as adequate representations of their separate course objectives, but that the partisans do not endorse the test as a whole as the proper embodiment of the course objectives because it includes objectives "that ought not to be included." In this case there is no blanket endorsement of the measuring device. The proponent of BDO then reiterates his claim that where there *is* such agreement, separate specification of what is to be learned is redundant. (As we shall see later, he may be wrong.)

Even the individual teacher concerned with the improvement of his own teaching effectiveness must measure that effectiveness by the evaluation of student performance. Whatever the objectives he sets out for student learning, he must be able to specify the sort of evidence that indicates whether or the extent to which those objectives are being met. If he cannot specify this, how can he hope to measure his effectiveness? If one cannot measure his effectiveness, how can he hope to improve it? If one cannot tell how effective he is as a teacher at any given time, then he cannot tell at a later time whether he is more or less effective than at some earlier time. If one is to improve his teaching effectiveness, then, he can do so intelligently or rationally only insofar as he can measure that effectiveness, only insofar as he can specify student performance as evidence.

Moreover, one must be able to specify the appropriate form of evidence with some foresight and consistency. The evidence gathered at different times for measuring his effectiveness must be comparable. To take a crude analogy, one cannot measure change in size over time if he measures height at one time and girth at a later time. Each is an appropriate measure of size, but they are different measures. One could define size as a composite of these two factors. Then he would need subsequent measures of both together to talk about change in size. If one is committed to the goal of increasing the size of certain objects and had chosen this way of measuring size, then he is committed to the objective of obtaining increases in the terminal height-girth composite. He has defined his objective in terms of his definition of size, but his definition of size is constrained by his desire to observe differences over time.

Similarly, if one were committed to improving his teaching of history, say, he would have to be able to tell how well his students had learned history under his instruction. He would have to decide what is to count as evidence of their having learned history, in a manner permitting comparisons before and after instruction. Measurements taken before instruction must be commensurable with those taken after instruction. To be committed to the improvement of one's teaching of the subject, then, would seem to mean to be committed to producing an increase in the difference between measurements taken before and after a specified teaching interval (whatever that interval might be). Whatever one means by history, he is, then, committed to the objectives of having his students improve their performance in coping with whatever devices he uses to measure their proficiency in history. However high-sounding the objective he would claim for the teaching of history, he can only justify claims for the *effectiveness* of his instruction to the extent that he can gather evidence appropriate to evaluating the impact of that instruction.

Good intentions do not justify claims for effectiveness. It is one thing to say what one wants to do; it is another to show that one has in fact done it. The rationale behind BDO recognizes this distinction and insists that if one is to show that he has in fact done what he has set out to do, he must first frame his intentions in such a way that he can tell when he has realized them. When one cannot have such evidence, he cannot tell whether he is improving the quality of his teaching effectiveness. It would seem to follow that BDO are essential for the improvement of teaching: there may be improvements in breadth, as the various aspects of subject matter learning are defined behaviorally; in clarity, as those aspects which cannot be thus defined are identified and accorded less attention in the curriculum; and in depth, as those elements which can be thus defined are afforded more attention and more opportunity is presented for their study.

Thus far we have been talking about BDO in rather positive terms. We have been putting forth the good intentions of those who would advocate BDO. If good intentions were enough to justify policy, we would not need to define educational objectives in behavioral terms. We could claim the good intentions implicit in our familiar garden-variety educational objectives as justifying educational business as usual, or, for that matter, as justifying any proposed innovation in schooling so long as one's intentions in proposing it were good ones. As the old saw has it, "The road to hell is paved with good intentions." But, then, perhaps so too is the road to heaven.

Good Intentions and Accountability for Endeavor

Effectiveness of instruction is not the only criterion for the evaluation of teaching or of curriculum. We are quite prepared to judge schooling for what it tries to do even though we are not at all sure how we should evaluate its success in so doing or even when the evaluation of success is determined otherwise than by examination of student behavior. We are quite prepared to complain that our schools are not trying to do something that they ought to be trying to do. We complain on the grounds that while we do not know how to account for their effectiveness in the particular endeavor we consider the endeavor important enough that they try to make a contribution on its behalf.

We might, for example, consider it most important that students be aided and encouraged to develop a sense of humanity, a concern for and with people, with mankind. It is not at all clear how much, if anything, schooling can do to influence such development. It may very well turn out that factors outside of schooling are the major determinants of this development. It seems likely that such development is highly dependent upon experience of the world and on maturity. It would be extremely difficult to measure success in the development of a sense of humanity while the child is in school. And yet, we may feel most strongly that we should promote this development in school as best we can. For some people, at least, schooling which did try to contribute toward the distant development of a sense of humanity would be an improvement over schooling which did not. Unless it would be shown clearly and unequivocally that schooling *cannot* make its contribution in that regard, the objective is too important, to some people, to be ignored.

This objective might be defined in terms of student behavior, but we have placed its attainment beyond the time (for many at any rate) when they are students. It is conceivable that we might agree on particular sorts of behavior of students while they are students which would serve as evidence that they are developing as desired. We should remember, though, that the effectiveness of our teaching is to be judged in the end on the fruition of the development of a sense of humanity in our students. To teach for the more immediate observable objectives presumes first, that we have in fact identified short-run behavioral correlates of a long-run behavioral goal, and second, that student behavior in school which might serve as evidence of the appropriate development is in fact more than role-playing on the part of students.

In the first case, for example, we may want to ask ourselves whether, say, the student's expressed and evidenced sense of fairness to others in his class or on the playground is evidence of his developing a sense of humanity. The answer is unclear. It may only indicate that he can care about those whom he knows or with whom he comes in contact. It may just as easily be a necessary condition of chauvinism as of global humanism. One might say that additional evidence is required to say for which of these one has better evidence. There are not many opportunities for students to show global concern for mankind.[4] (For that matter there are few opportunities for most adults to show such concern. Still, we consider it important that adults be prepared to show such concern under appropriate circumstances.) If we wish to create opportunities for students to exhibit behavior which would give us appropriate evidence and if we are to define objectives in terms of such behavior, we must be careful to avoid artificiality. Thus we might wonder about evidence gathered in a written exam[5] in which the student exhibits the behavioral objectives by answering certain questions in particular ways. (This does seem to be stretching the idea of behavioral objectives pretty far, but then the notion is ambiguous enough to be stretched that far. More on this point later on.) Surely we would not want to contrive situations in which some students are put in need of compassion and assistance. We might try to have students help those who are disadvantaged in fact, in hopes that thereby they will exhibit the appropriate behavior. Under this circumstance they might well exhibit the behavior but they might not be developing the sense of humanity.

---

[4] Some of my students have suggested the antiwar mobilization rallies as a counterexample. But, first, these are rather rare occurrences if one takes the perspective of, say, twenty-five years or more. Second, the motives of those who engage in them are mixed and it is hard to tell whether one is witnessing a major outpouring of global humanism or a combination of international isolationism and nationalistic chauvinism.

[5] It is conceivable that people answer as they might act or that having been required to think enough to answer questions on such an exam, their subsequent actions might be appropriately influenced. Still, even the authors of the most extensive studies involving written examinations in the moral and humanistic spheres were careful not to claim too much. See Hugh Hartshorne, Mark A. May, et al., *Testing the Knowledge of Right and Wrong*, The Religious Education Association, Monograph no. 1, July 1927, especially pp. 1-2: "One of our problems is to discover what the relation *is* between behavior and the knowledge of right and wrong. . . . It may be that overt action is a far better indication of what a man really knows about right and wrong than his verbal responses are. . . . [However] one can find out by word responses whether an individual is aware of certain customs even though his possession of a certain custom in the form of a habit is not revealed."

They might be role-playing. Students are often quite adept at exhibiting the sorts of behavior and giving evidence of the sorts of attitudes that are expected of them in school.[6] It should not be too surprising if they are also adept at exhibiting the sorts of behavior and attitudes expected of them outside of school. If these environments have contrary expectations, the "well-adjusted" student may be the one who meets the evidential expectations of each environment when he is in it, thus leaving open the question of the state of his own attitudes and the direction of their future development. Whether the evidence is called for in class discussion or on a questionnaire, the student learns to give the "right" answer. If helping those underprivileged folks from the ghetto is the "done thing," the student does it. Whether his doing it and even doing it with energy, imagination, and sensitivity means anything more to the development of a sense of humanity than being a member of the drill team would have meant is open to question.

Is it open to empirical question? It is possible that questionnaires might be designed to survey attitudes of adults on humanism. Then one might try to discover whether and how educational experiences related to the results of the questionnaire. If the questionnaire is, in fact, well designed to test for the appropriate attitudes, then it is an empirical matter of determining whether and what sorts of educational experiences relate to the various findings. The three sorts of interesting results of such an investigation for our immediate purposes might be:

---

[6] It may, of course, turn out that getting students to play certain desirable roles is the best hope we can have of recruiting them to such roles "for real."

1. Adult attitudes toward humanism vary inversely with educational experiences.[7]
2. There is no significant relationship.
3. There is a clear positive relationship.

If the first occurred, we would have a fairly strong case for avoidance of educational efforts in behalf of the development of humanistic attitudes. I am willing to bet that it would not turn out thus.

What of the second case? To say here that there is no significant relationship is to say that there is fairly equal likelihood that people with divergent attitudes as adults had similar educational backgrounds as far as development of the attitude is concerned. It may very well be that the educational objective is so popular that all receive some such educational treatment, but with diverse developmental results. Perhaps other factors might be identified to account for the differential results, thus reducing the diversity in findings. On the other hand, we might notice that rather than commonality in relevant educational background, there is diveristy in educational background which cuts across the attitudinal results. People from diverse educational backgrounds are represented all along the whole continuum of responses on the attitude survey in much the same proportions. Then, in the absence of further information, we might want to argue that there is no clear evidence that the educational background helps or hurts in the development of presumably desirable attitudes, but that as long as we do not know clearly that it hurts, we should see if we can get it to help. We can do something that might be innocuous out of good intentions. Of course, it is not clear in which direction the balance will be tipped by our good intentions.

The third possibility seems on balance to suggest that because of past success we should continue to educate in hopes of influencing development of a sense of humanism. Again we raise the issue of role-playing. Does the fact that people who have been educated to respond to particular attitude-sampling cues in particular ways tell us very much about their attitudes? As Myrdal[8] pointed out a long time ago, we are quite prepared to express attitudes of global humanism while living in accordance with chauvinistic principles which are in clear contradiction of humanism. We learned in school that all men are created equal, but how many of us have behaved consistently as though we believed it?

---

[7] We are treating all such experiences together. Clearly one would want to distinguish among various sorts in any sophisticated study.

[8] Gunnar, Myrdal, *The American Dilemma* (New York: Harper & Row, 1944).

We might, then, insist on the empirical test of behavior of adults as evidence of attitude. Such evidence may be extremely difficult to gather to any significant extent for large numbers of people. Most of us live a good part of our lives isolated or insulated from circumstances which put such attitudes to the behavioral test. When they are evoked in behavioral situations, the situations are often "contaminated" by other considerations as, for example, good public relations, relations with fellow employees or with the neighbors, or with the law of the land. Surely we would not want to base our judgments on the exceptional cases—the zealots or the martyrs. We might, then, conclude that expressed attitudes are the best we have to go on and concentrate our efforts at developing better and better devices for measuring attitudes, perhaps through checking on consistency in expression of attitudes on a given survey instrument. We might take the position that it is reasonable to hope that those who express appropriate attitudes are likely to live by them or are susceptible to moral suasion on being reminded of them, and that if some educational endeavors are consistent with subsequent expression of appropriate attitudes, then those endeavors are worthwhile.

But someone might remark that this is tantamount to the justification of the objective of developing a sense of humanism in behavioral terms. If it is, then BDO are hopelessly, ambiguously trivial. For it is not clear that the behavior expected in the student as a student must be the same as that elicited from him as an adult. More to the point, if the behavior expected of him as a student and/or as an adult is the expression of appropriate verbal responses to particular sorts of cues, then mere rote verbalism is behavior. In that case there is no distinction in BDO between parroting and conviction, between responding with as against responding without understanding of the response.[9] One wants more of the notion of behavioral definition of objectives than the idea that so long as the student does something and that something is specified by the teacher as an objective, then it is a behaviorally defined objective. All one need do is to begin every currently nonsensical or ambiguous educational objective with an expression such as "the student says . . ." or "the student reads . . ." and it is magically transformed into a behavioral objective.[10]

---

[9] See Israel Scheffler, *Conditions of Knowledge: An Introduction to Epistemology and Education* (Chicago: Scott, Foresman & Co., 1965), pp. 76-79.

How much better, more intelligent, more honest to say that an objective such as the development of a sense of humanism is to be adopted because it is worthy of being fostered. We do not know for sure how to measure our effectiveness in attaining this objective, though we can, as suggested, try to develop greater sophistication regarding it. Still, we might want to say that a curriculum in which people are encouraged or required to try to make a contribution is better than one in which people are not thus encouraged or required or one in which the objective is ignored. Before our enthusiasm gets the better of us, we might attend to that nagging suspicion that it is one thing to have good intentions; it is another thing to decide among alternative ways of pursuing these intentions. It is yet another thing to come up with a rationale justifying the amount of time and effort in schooling that should be devoted to alternative good intentions, especially when some are more specifiable in concrete terms than others. But there are still other sorts of teaching intentions to discuss.

Good Intentions—Once Removed

At the beginning of the previous section we suggested the possibility that *success* of schooling might be evaluated without regard for student behavior. Schooling might be aimed at preparing the student to do his own thing if and as he can and wishes. The emphasis is on preparing and "preparing" might be construed more or less loosely. At one extreme it could mean rigorously fitting with a variety of skills and an abundance of knowledge. In that case it could just as well be handled under the heading of the usual sorts of accountability notions, those in which the schools are held accountable for some measure, at best, of what the student does or can do with his life. We are more interested here in the other end of the continuum, the end where the notion of successful schooling is harder to grasp, if indeed it is there to grasp at all.

---

[10] Mager, *Preparing Educational Objectives.* This book abounds with appropriate examples.

If a teacher claims only to present a set of alternative interest lines, or career lines, or study lines to a student, say, we can look to see if the set was presented. Was it intelligible? Was it complete? Were the students aware that it was being presented? We might say then that the set of alternatives was successfully presented. We need not care whether the students took much note of the set, asked questions regarding it, or did anything at all as a result of having the set presented to them. Analogously, a teacher may present a series of lectures, discussion sessions, and work assignments which he feels might appeal to those interested in history, and his test of whether they are interested in history is whether they take advantage of the offerings and *learn* history from what he has to offer. (This is not a bad description of much teaching, I would bet.) Success then is judged simply by whether the series was offered. If there were no takers, we might wonder whether offering the series requires more than merely announcing it, or would the teacher have to lecture in an empty room and also hold discussions with himself? This might sound silly but college departmental course requirements are often more readily explained by reference to such considerations than by the official reasons given for them. At any rate, there is a claim to success in teaching under these two examples based on intention to present, and it is not even clear that the presentation must occur.[11]

Aside from the fact that we boggle at the idea of successfully intending something—this does seem to be a bit of semantic sleight of hand—some of us are likely to be bothered by the lack of responsibility for the success of our teaching in the rather traditional sense of concern that our students learn what it is that we would try to teach them. At a time when there is a general sense of dismay at the results of so many well-intentioned public policies and public programs, at a time when a good number of students question the wisdom or relevance of the content of their courses and the content of the curriculum as a whole in their schools and colleges, it is not surprising that a good many faculty have chosen to let students do their own thing. Only they, too, choose to do *their own thing as well.* Often that means to continue to offer whatever they have been offering in whatever ways they have been offering it without having to worry any more about whether students are learning anything from them. Those who worry about loneliness for one reason or another might alter their courses in ways they calculate will generate some student following. If they are still seriously worried about their students learning from them what they have to teach in these new courses, then they do not belong under this heading.

---

[11] See Israel Scheffler's *The Language of Education* (Springfield: Charles C Thomas, 1960), chap. V, "Teaching and Telling."

There is some room in this category for those whose good intentions are bound up in a humanitarian concern for the emotional and mental well-being of the student. There is a concern for the development of the student's identity, his sense of self, a healthy self-concept. He must learn who he is *really*. He must be treated as a person. He must be accepted as a person. He must develop in a manner consistent with the springs of his own inner development. It is not clear what he should learn if all this is to happen, nor is it clear whether the teacher who professes such notions as these approaches his students with anything more than "Do your own thing." If he does come with more than this, if he does have expectations regarding what he should teach and what his students should learn, then he comes to teaching with more than his rationale or his rhetoric will justify.

Now, it may be that he does have some things he would teach, if and as his students see themselves as unfolding in the direction of learning it. But he avoids responsibility for deciding what students ought to learn.[12] He might plan nothing but asks his students what might interest them. He might put all sorts of interest catchers in their way and wait to see what catches their interests. He takes a sort of *laissez faire* Divine-Hand-of-Providence view of the world. In due time they will seek to learn what they need to learn. At any rate, whatever they seek to learn may be presumed to be what they need to learn as the teacher helps them to positive self-concepts in which they very well might see themselves as happy. Or again, *realistic* self-concepts might require some of them to see themselves as ignorant. Conceivably, positive, realistic self-concepts might lead some of them to make true the belief that they are happily ignorant.

---

[12] Some of those who talk this way may still presume to know just what each of his charges needs to develop his own unique self.

Just as it is easy to fall into this sort of overstated or misstated conception of respect for students as persons, it is easy to presume that the idea of taking responsibility for the effectiveness of teaching is somehow inconsistent with respect for students as persons. It is hard to imagine in this day and age that students should not learn to read. Learning to read involves a good deal of what can only be described as sheer drudgery. Respect for the persons that our students *might become* leads us to judge schooling on its effectiveness in teaching reading. We may try to reduce the drudgery, make games of much of it and engender enthusiasm to undergo what must be undergone to enjoy the rewards of literacy. Still, few of us, I would bet, are so willing to respect the student's right to do his own thing as to permit him to remain for long an illiterate. The test of his literacy is his demonstrated competence in reading. The test of our concern for him as a person is in how far we are willing to go to keep him from illiteracy and the unfortunate consequences thereof. Moreover, one may treat the student in a most humane and respectful manner while working at the business of teaching him to read. The idea that there is an inherent contradiction between humanistic concern for students as persons and the behavioral definition of objectives is absurd. It leaves plenty of room for the collision of these two concerns, however.

Loaded Dice

Empirical research on teaching and learning is, as has already been indicated, aimed at evaluation of the success of teaching. Success is translated into effectiveness; teaching intentions are translated into intended outcomes of observable behavior (or outcomes inferable from observable student behavior).[13] How, then, are we to undertake a comparison between thus defining objectives and not doing so? Aren't the dice loaded?

---

[13] For those who would like a more direct examination of the notions of behavior and behaviorism, see my discussion of those topics in *Educational Myths and Realities: Philosophical Essays on Education, Politics, and the Science of Behavior* (Reading: Addison Wesley Publishing Co., 1958), Part II, "Behavior, Behaviorism and Teaching." I attempt to show, hopefully in a useful way, just how ambiguous those notions are. The ambiguity of such

Normally we take teaching as an enterprise aimed at learning. Teachers are expected to teach what students are expected to learn. It may turn out that our expectations are unfulfilled from time to time. We feel justified in evaluating teaching on the basis of both sorts of expectations. We may ask whether teachers are teaching what the students are expected to learn, and we may ask whether students are learning what it is they are expected to learn. On the one hand we may be most interested in examining teacher intent. On the other hand, we may be primarily concerned in evaluating the teacher's success in carrying through the intent.

Still, we should not forget the relationship of these expectations. Students may learn all sorts of things that they were not expected to learn, i.e., that were not part of the teacher's intent. Some of these things might be quite useful to students as, for example, how to survive in school with a minimal amount of effort and aggravation. If the teacher had not intended such learning, then we should hardly credit him with teaching success for it, even though it may have been learned more effectively in his class than in any other. We should say here that success in teaching presumes appropriate intentions by which to judge success.

Not only may teachers do things which lead consistently to learning which they never intend—teachers may claim to intend learning which does not or cannot occur, or learning the occurrence of which cannot be judged. If teachers are expected to teach what children are expected to learn, then we may ask whether children may reasonably be expected to learn what their teachers claim to intend that they learn. Thus, for example, suppose a teacher claims to intend that her students learn some specific skill and they fail consistently to develop that skill. We should wonder whether there was something about the students that precluded their de-

---

notions as behavior and behaviorism does not, however, justify the sort of treatment they get in French et al., *Educational Goals*, and in D. R. Krathwohl et al., *Taxonomy of Educational Objectives: Handbook II: Affective Domain* (New York: David McKay Co., 1956), in which the word *behavior* and the word *objective* are often used interchangeably (more so in French), thus rendering vacuous the notion of behavioral objective.

velopment of the skill and ask, "Was it reasonable to expect *them* to develop the skill?" We might then talk about such things as characteristics of maturity or requisites of prior learning or experience. Or, we might focus more directly on the teacher and ask, "Was it reasonable to *expect* the students to develop the skill?", which might be translated in this case as "Was the teacher's claimed intention genuine? Did the teacher actually try to get them to develop the skill?" We might then talk about empty promises or about ritualistically going through the motions of teaching routines that do not teach.

Suppose, however, that the teacher claims that his teaching is aimed at the student's development of some character trait or attitude as, for example, a healthy, positive self-concept. Or, suppose that the teacher would try to contribute to the child's "sense of oneness with all things" or to his sense of humanity or to his development as an independent, mature, happy, socially responsible adult. In these cases it is not at all clear just what the teacher is expected to teach; it is not at all clear just what the student is expected to learn. There is a lack of specificity in the meanings of such objectives. The "sense of oneness in all things" sounds impressive but what on earth does it mean? All things that can be distinguished as individual things have at least that characteristic in common; but do they have any other characteristics in common worth noting—worth noting in ways that physical science does not note? We should probably get into spiritual or spiritistic interpretations of the expression and find ourselves embroiled in disputations that do not lend themselves well to resolution by reference to empirical considerations.

How, then, is the poor teacher or researcher to evaluate the effectiveness of teaching for such an objective? Is it fair to insist on the evaluation of *teaching effectiveness* for such an objective? To the extent that such an expression is not amenable to specification in BDO terms, the empirically oriented teacher would be likely to rule it out of the curriculum in favor of those hard or substantive areas that can be so specified. The more spiritually oriented would be likely to resent the BDO pushers for so doing. Matters are not likely to be helped by an attempted rapprochement whereby the empirically oriented would offer to tidy things up by specifying the expression in BDO terms.[14] The spiritually oriented may insist on the sanctity of spirituality. How are we to choose between spiritually rationalized and BDO-rationalized components of the curriculum if we have no meeting ground for their rational comparison?

And we might ask: What is a properly developed balance of independence and social responsiveness for a mature adult? How independent should a man be? How responsive? Which is more important, independence or responsiveness? With what should such an adult not be happy? What does a sense of humanity require on a continuum between complacent fellow feeling and total commitment to the betterment of the circumstances of others? [15] What is it that students are expected to learn; what is it that teachers are expected to teach?—answers to these questions?

In part, that must depend upon whether there *are* answers to these questions. The problem is that there may be several sorts of answers. The issues involved are issues of cultural relevance and in a large society such as ours they are issues of subcultural relativism. The problem of specification is complex. One must try to get at what the members of various subcultures hold to be correct answers. One must face the philosophical question of whether there ought to be rational agreement on particular answers across subcultural lines. One must face the related question of the political necessity for agreement on answers to such questions as these. And, one must face the difficult problem of relating professed or verbal attitudes on all these issues with their behavioral correlates.

---

[14] It is conceivable that one might find evidence that someone believes any sort of ephemeral or even nonsensical idea. It is conceivable that one might take as his objective that the student should believe that the moon is made of green cheese, and that one could define this objective behaviorally, i.e., in terms of what sort of evidence would count as the student believing that the moon is made of green cheese. Still, where it is not clear what sorts of things a person *does* who believes in, say, the universal spirit in all things, then the very specification of behavioral objectives defines what is to count as believing in that spirit and thereby defines belief in that spirit. This is the sort of problem that tends to get swept under the rug in Krathwohl's *Taxonomy* . . . despite caveats to the contrary. See especially 5.2 in *Handbook II* on "Characterization," pp. 170-75, 185, where the authors get to the level of internalization of philosophy of life.

[15] These are the sorts of issues swept under the rug in French et al. in the attempt to objectify pious platitudes so as to square concern for social welfare and the political interdependence of mankind with the ethos of individual social and economic responsibility. Here are a few sample "Illustrative Behaviors" of the ideal product of general secondary education: (a) "He considers a vocation from the standpoint of service to his fellowmen as well as from its financial possibilities," p. 190; (b) under 2.232, "Recognizes that no right is absolute and that all rights are limited when they interfere with other rights"; and (c) "Recognizes the rights of individuals to own, control and use private property," p. 152.

For it is one thing to say these must be verbal agreement on the answers and another to say that people must evaluate behavior by the answers to these questions in a consistent fashion. Thus, for example, we might agree on the idea that responsiveness to opinions of and concern for others should take precedence over independence, but disagree as to when a given act shows independence rather than insensitivity to others. It is conceivable, then, that we might agree on a host of objectives at the verbal level, yet disagree at the behavioral level. It is also possible that we would evaluate behavior in the same way though disagreeing verbally. This is the case of "mere verbal disagreement."

In either case we would agree, I think, that our evaluation of behavior is the ultimate test of the sorts of attitudes we have been talking about. This is still not a naive call for behavioral definition. For insofar as we might evaluate various behaviors differently, behavioral definition has not solved our specification problem. It merely states it more clearly. That may be a gain in and of itself. Nevertheless, we cannot proceed directly from the statement of potential disagreement in behavioral definition to the specification of BDO for education. His behaviorism does not give the educational researcher or objective-setter the right to choose among alternative behavioral specifications of a concept if in fact there are alternatives.

It is of equal importance to point out that even when the alternatives appear to be the adoption of an educational objective in non-BDO terms as against one in BDO terms, behaviorism alone is not sufficient to justify the choice. At the very least, the behaviorist should be aware of the possibility of alternative behavioral specifications of the objective in question and not select the first one that suggests itself to him as *the* BDO. As indicated earlier it is much too easy to put any silly notion into BDO form by using an activity verb. Thus, for example, if "understands modern poetry" is not BDO, one changes it to "writes critical essay on modern poetry" and all is well.

It is too easy to put a very important notion in BDO terms, thus rendering it confused or misleading. To return to an earlier example, we remark that it seems worthwhile to take as an objective an understanding of mathematics—say, in particular, algebra. Now doing algebra is not some type of behavior even if it is an activity. Solving equations of the . . . sort is not the whole of doing algebra either, no matter how exhaustive the list, if doing algebra is to be correlated with the notion of understanding mathematics. For we would expect in that case that one who understands algebra would know not merely how to solve a group of equations but also when one sort rather than another is called for—i.e., we would expect him to be able to use some mathematical judgment. And, while one may show judgment in the way he does things, judgment is not behavior.

The very same empirically observable behavior may signal mathematical judgment or rote learning, depending upon the sorts of teaching and experience that precedes the exercise of that behavior. The problems on a given algebra test may be new to one group of students and be the very ones on which another group has been drilled. The same test, then, may require the same *observable* behavior—solving identical problems—as evidence of the success of radically different teaching objectives. If the test *is taken as* the behavioral definition of each of these different sorts of teaching objectives, then BDO cannot distinguish between the objective of having students develop proficiency in the exercise of mathematical judgment in doing algebra and having students reproduce solutions to specific types of problems on the basis of rote learning. Again, the same examination may be appropriate to both objectives, but only if the objectives are specified by additional information can we say that the exam enters into the behavioral definition of objectives. Judgment is not behavior; it is inferred from behavior and appropriate background. Analogously, one may do things in a manner showing fine aesthetic sensitivity or intelligence or creative imagination, but these are not themselves behavior.

One may take as his educational objectives the development of whatever potential his students might have in these respects, and he may indicate what kinds of activities he intends that they engage in order to develop this potential, but has he defined his objectives in behavioral terms? Yes and no. If a teacher claims that he wants to develop these attributes in his students and yet he could give *no examples* of the sorts of things that his students might do by the end of the course which would give evidence of such development, then we might legitimately wonder whether the teacher's claimed intentions were serious or even rational. We would not accept the specification of a list of activities and their components to be gone through mechanically as behavioral definition of an objective purporting to represent the development and exercise of creative imagination. But we might at least be willing to accept the idea of behaviorally defining objectives as the provision of a set of examples *of the sorts* of things we might expect students to do (with an indication of the kinds of backgrounds which they might bring to doing them) if they had developed in the hoped-for manner. One does not have a routine for imagining creatively. Still, if we are seriously concerned about the development of judgment, sensitivity, and creative imagination, we must be careful not to dismiss BDO as a caricature only to adopt the equally sterile caricature of irresponsible antiempiricism.

Empiricism and the Sensitive Educator

All our judgments about teaching are not to be made in terms of effectiveness alone. There are, as we have suggested, some ways of treating people that are better than others, if only because they show a greater respect for people as people. This does not provide a license to presume that children always know what is best for their own development as persons, but it is a healthy corrective to the often observed treatment of children as though they were nonpersons. The current emphasis on such concern for personhood is likely to show itself in a renewed hostility toward empirical research in the field of education. Scholarly or clinical detachment of the researcher in behalf of the objectivity of his research will be translated in the minds of some into disinterest and lack of concern for the personhood of children. This sort of attitude is a sadly mistaken one.

Although some people will show diminished interest or hostility toward research in programmed instruction, teaching techniques, student motivation, and curriculum development, advancement in research in all these areas would be consistent with the idea of helping students to "do their own thing." Curriculum development especially might need to give considerable attention to the problem of including an increased number of alternatives at various levels of interest and ability. Programmed instruction may be viewed as a device for expanding curricular opportunities. Teaching techniques that are more responsive to divergent levels of student interest and ability may need to be developed. Or, perhaps more honestly, we may need to develop techniques by which teachers may cope with such divergency while maintaining harmony and without destroying interests of students. Motivation studies are called for at least to identify those factors which impede student motivation. There may be a stronger interest in mental and emotional development. There may be renewed interest in the relationships among personality, interest, life-style, and occupational choice. It might even be possible to mount significant empirical studies to determine whether children do come to want to learn what they need to learn and, if so, what factors operate to influence them in appropriate directions. Of course, there are some tricky questions about determination of "need" and "appropriate" which are not wholly empirical in nature.

At any rate, it seems clear, even to the most antagonistic opponent of behaviorally defined objectives, that there is quite a bit of empirical work of ultimate value to be done in educational research. Hopefully, "behaviorism" is ambiguous enough to absorb the antagonistic and to permit the advancement of empirical research in education.

S.C.E.R.T., West
Library
Calcutta
25/3, B C. Road

CHAPTER VII

Individualized Instruction: A Projection

JAMES E. MCCLELLAN

To venture yet another essay on the topic of individualized instruction may well appear an act of hubris. In the literature of educational exhortation no less than in the literature of educational research, individualized instruction (II) is a perennial favorite. Is there anything about II which has not already been said many times by psychologists (of every ilk), sociologists, geneticists, classroom teachers, and students, and, yes, philosophers of education? I believe, of course, that there is something new and important to be said, else I should not present this new attempt. But let me first review some of the things which have been said.

A good place to begin is in the first half of the decade just preceding this one. Many things were happening to schools which occasioned another look at this, by now hoary, topic. Two of these developments are worth particular attention—the coming of age of programmed instruction (PI) and the new awareness of the deprivation of many groups of school children, especially black and Puerto Rican children in urban slums. I was involved in both those movements in a somewhat peripheral way, and I should like to describe them in an autobiographical mode of speech. But before doing that, I should beg indulgence to establish an informal base line to show approximately where thought about II stood prior to the impact of PI and the War on Poverty.

I

There is a wonderful thing about yearbooks—they come with dates attached. The NSSE yearbook (Part I) of 1962 was entitled

Individualizing Instruction.[1] The ASCD yearbook of 1964 bore the identical title.[2] There is otherwise little connection between the two works. The ASCD yearbook committee, under the chairmanship of Ronald C. Doll, must have been well along in its collaborative effort before the NSSE volume was published; the one reference to the earlier work in the later one seems an afterthought. One might complain that the isolation of these two efforts, each an official output from a great national organization, represents a deplorable fragmentation within the educational establishment of this country. The fragmentation is certainly real enough, but it would have been more deplorable had anyone possessed the power to cram such disparate works within the same cover. It is their contrasting treatment of the same topic which warrants our continued interest in these two volumes.

The NSSE yearbook, as befits its Oxford grey and gold cover, is perfectly straightforward in its approach to the nature of individual differences. People are different; so much is patent. What now needs study (The word Scientific has been dropped from the Society's name, but its impetus is still felt) is exactly how and how much and how come school learners are different. The chapters march in ordered column: Differences among learners arise from the genes. Differences arise from social variables. Differences among learners are magnified or reduced by factors within schools themselves—organization, teachers, curriculum, and so forth. Chapters on each. And so on, all very rational and sequential, like a well-ordered school itself.

Three major conclusions emerge from a careful reading and rereading of this NSSE yearbook.

First conclusion: Careful and methodical research will always reveal a range of variation much greater than we would have believed on the basis of casual observation alone. This consistent underestimation of variation holds true for physical as well as mental or psychological measures. One naturally formulates hypotheses to explain why we tend to see children as more alike than they really are. But "why" is not the question here. The yearbook can be seen as one good answer to the question: What have we learned from a half-century of educational research that we didn't already know? The more carefully and methodically we carry on research, the more we are forced to recognize that school-age children vary enormously on just about any scale we choose to apply, vary much more than we would have believed.[3]

---

[1] *Individualizing Instruction*, Sixty-first Yearbook of the National Society for the Study of Education, Part I (Chicago: Distributed by the University of Chicago Press, 1962).

[2] Ronald C. Doll et al., *Individualizing Instruction*, 1964 Yearbook of the Association for Supervision and Curriculum Development (Washington: the Association, 1964).

Second conclusion: The first conclusion creates a peculiar ambivalence among responsible educators. On the one hand, they are ideologically, perhaps sentimentally, committed to individualism (Beck's essay dealt with the European sources in that fairly confused ideology), but on the other hand they are quite well aware that excessive variations among pupils in a class seriously disrupt effective instruction under normal school procedures. Quite often in the yearbook, one reads an author who gives sound, practical advice on "how to take care of" individual differences in the classroom, and then the same author will wholeheartedly affirm: *Vive la différence!*

---

[3] I am indebted to Dr. W. Timothy Weaver, Research Fellow in the Educational Policy Research Center, Syracuse University, for providing me the bibliographical references used in this essay and for correcting many of my misconceptions about how research in this area is actually carried on. Dr. Weaver caused me to question the persistent claim in the NSSE yearbook that we underestimate variability in children. Lee J. Cronbach writes: "It is very likely that teachers overdifferentiate. I know no research on impressionistic adaptation of instruction, but something can be learned from studies in which counselors have been asked to predict a student's grade average. Various biases appear in the estimates, but the most significant finding is that the counselors overdifferentiate; they tend to expect too much from the persons who tested high, and too little from those who tested low. A regression line is an actuarial formula that starts with the group average and ranks differential information as a correction factor, giving the latter just as much weight as it deserves. The greater the accuracy and pertinence of the differential information, the greater its weight. The judges in the study cited gave considerably more weight to differential information than the regression formula did. Certain reasonable assumptions, entered into a decision-theoretic model, lead us to the conclusion that the poorer the differential information, the less the teacher should depart from the treatment that works best on the average. Modifying treatments too much produces a worse result than treating everyone alike." See Lee J. Cronbach, "How Can Instruction Be Adapted to Individual Differences," in *Learning and Individual Differences*, ed., Robert Gagné (Columbus, Ohio: C. E. Merrill & Co., 1967), pp. 29-30.

Perhaps these two positions—Cronbach's and the yearbook's—are ultimately compatible. Perhaps counselors and teachers underestimate the range of variation among youngsters and still overestimate how much to modify treatment (therapy or instruction) to "take care of" individual differences. This double miss seems a likely consequence of a situation in which youngsters are undergoing instruction by compulsion rather than by choice.

> "Much of the delight of social living springs from the differences among individuals. Also deviation is at the base of creativity, so much needed in every age and certainly in ours."[4]

What is a bit peculiar is that the authors seem less than aware that they *are* ambivalent. They seem to believe that their ceremonial-like affirmations of fealty to individualism express their true values, while their remarks on "how to handle" individual differences in the classroom are merely technical advice. Educators who affirm the value of individuality and then give advice on how to handle individual differences in a conventional school setting have a more heroic role than they seemed (in 1962) ready to play.

Third conclusion: Research has proved the ineffectiveness of certain standard devices for "handling" individual differences administratively, but the advice given to teachers is to use these same standard devices in classroom pedagogy. But these devices will continue to fail wherever applied. What explains this puzzle?

The three standard administrative devices for handling individual differences are retardation of slow students, acceleration of the bright, and enrichment for the bored. It's a fascinating incident in cultural history that we spent an awful lot of time and energy to demonstrate empirically what could easily be proved by reasoning alone. To retard slow students will widen, not reduce, the range of variation at every grade level except the highest. It may not be as obvious that accelerating able students will have the same effect except at the lowest grade, but a couple of minutes' thought should make it obvious. And enrichment, which can be justified on purely humanitarian grounds (Why shouldn't those who are bored with the instructional program have some other harmless activity available?), has no effect at all on the range of variation in classrooms.

These remarks apply in the first instance to elementary schools, where children are grouped on some basis or other in the first grade and are promoted, skipped ahead, or left behind at the end of each year. I now want to affirm that the basic problem in homogeneous grouping arises wherever one attempts to handle individual differences by reducing the range of differences in the classroom. That is what ties together the topics of individual differences and homogeneous grouping. If a teacher is giving instruction in arithmetic to a group which varies quite markedly in arithmetical competence, then that teacher is going to have great difficulty in providing effective instruction to the group as a whole; other techniques, probably quite expensive and possibly distracting, will have to be employed if instruction is to be given effectively to each individual pupil. One can try to reduce the range of variation in the instructional group to the point that instruction given to the group as a whole will be entirely appropriate to each individual in the group. Thus we may increase the effectiveness of instruction without excessive costs. That, at least, is the conventional wisdom, and the empirical studies reported in the NSSE yearbook do not seriously threaten it, though they do make it quite clear that the conventional wisdom is not a panacea for pedagogical problems arising from individual differences.

---

[4] Esther J. Swenson, "Teacher Preparation," in *Individualizing Instruction*, chapter XV, NSSE Yearbook, p. 296.

One problem that the conventional wisdom has trouble with is keeping the homogeneous groups homogeneous long enough for instruction to remain effective. Given normal school procedures, including that of giving instruction to groups as wholes, a necessary condition for effective instruction to individuals is a reasonably narrow range of variation in each group. I shall call this last sentence, with its qualifications, Principle NRV (for narrow range of variation). Given normal school procedures, including more educationally relevant materials available than any one student will use, a rather open-ended set of objectives for any particular instructional activity, and so on, a sufficient condition for increasing the range of variation (among the members of the instructional group in progress toward that open-ended set of objectives) is effective instruction itself. I shall call this last sentence, with all its qualifications, Principle EI (for effective instruction). Now principles NRV and EI do not contradict each other; they are both true, or so I believe. But taken together they mean that any attempt to handle the problem of individual differences by establishing reasonably homogeneous groups (in accordance with NRV) at time, t, is going to fail as a solution (in accordance with EI) at some subsequent time, t_2. The only sure ways to keep EI from operating is to reduce class size to one member or the length of time a class of two or more students stay together to the time it takes to perform one instructional act and then rehomogenize.[5]

I am not exactly sure of the status of Principle EI in this argument. It stands to reason, as we might put it, that effective instruction should widen the range of variation. Think of a basic literacy program and imagine a group of a hundred adults, all totally innocent of the basic idea behind reading, i.e., the idea that language can be represented by visually perceived symbols. In that group the range of variation in reading skill is nonexistent. Effective instruction, we would expect, would widen the range of variation until at some point the group would have to be subdivided if competent instruction were to continue.

Principle EI not only stands to reason, it is confidently affirmed by the research specialists in the field of individual differences. Indeed, EI is taken so much for granted that some authors appear to assert that measures of the range of variation are measures of the effectiveness of instruction.[6] Other authors declare that if teaching advances the mean of achievement in a group of students but does not widen the range of variation, then the objectives of the teaching are too narrowly defined. From that point of view, then, there is a powerful moral argument that teaching *ought to try* to widen the range of variation.[7]

---

[5] A variation on this second technique is to establish for any instructional group a very limited set of objectives which must be mastered by every student in the group, and then to promote any student out of the instructional group when he masters those objectives. This strategem is not discussed in the NSSE yearbook either because it was invented after 1962 or because none of the authors found it worth recommending.

[6] Walter W. Cook and Theodore Clymer, "Acceleration and Retardation," *Individualizing Instruction*, NSSE yearbook, esp. p. 191. There is a great deal of highly technical argument on this point. If we distinguish practice from instruction, we may appeal to a body of research on the effect of practice. By 1925 it could be concluded "that experimental studies yield varying answers to the question whether individual differences increase or decrease as a result of the addition of equal amounts of practice. The preponderance of evidence leads to the conclusion that differences generally decrease, leaving things inconclusive." (Robert Glaser, "Implications of Previous Work on Learning and Individual Differences," in *Learning and Individual Differences*, pp. 2-3.) This same inconclusiveness carries through all the more recent studies of the same phenomena. Perhaps this is *why* various authors in the NSSE yearbook insist that instruction *must* increase the range of variation: maturation and practice (together or singly) may increase the range of variation or they may reduce the range, but to have given effective and individualized instruction *implies* that the range of variation increases; i.e., it is a *conceptual* distinction.

---

[7] "It is a part of the educational folklore that there is an increase in the variance of the achievement scores of students of the good teacher, since he stimulates bright students to forge ahead to a greater degree than does the mediocre teacher. The popular answer of persons in programed instruction is no . . . with effective instruction even the dull student will achieve a high level of performance. The fact of the matter is that in and for itself . . . [an increase in variance in achievement] is neither favorable nor unfavorable. It all depends on the *means*. When the mean level of achievement under a certain training condition is higher than the level under a standard condition, that is good; when the mean under the training condition is lower, that is bad." (Richard C. Anderson, "Individual Differences and Problem Solving," in *Learning and Individual Differences*, pp. 81-82.

Anderson's point might have been stronger had he intended the other sense of "*means*" (his emphasis). It depends more on *how* you do it than on whether or not you raise the average achievement.

I am inclined to believe, however, that Anderson's summary does not entirely do justice to the proponents of programmed instruction. The sophisticated position there, I think, is that of Yeager and Lindvall: ". . . one may say that these results suggest that level of attainment reached prior to entering IPI is correlated with IQ, but the actual rate of progress during one year of Individually Prescribed Instruction is not correlated with IQ." (John L. Yeager and C. N. Lindvall, "An Exploratory Investigation of Selected Measures of Rate of Learning," *Journal of Experimental Education* 36 (1967): 79. (Also as University of Pittsburgh Learning Research and Development Center Reprint 28.) The publications of the LR&DC are an unrivaled source for understanding the current research emphases in individual instruction, IPI being the LR&DC's own version of a system approach to the use of PI and other techniques in working with individual learners. Cf. C. M. Lindvall and John O. Bolvin, "Programed Instruction in the Schools: An Application of Programing Principles in 'Individually Prescribed Instruction'" in *Programed Instruction*, NSSE 66th Yearbook, Part II, 1967, pp. 217-54.

The *moral* emphasis underlying IPI, it seems to me, is *not* the empirically false claim that the dull student will achieve at the same rate of competence as the bright student, but rather that if we are really attentive to the range of tasks we include in an instructional program, we will find that we have to give up our notion of students being bright or dull *überhaupt* and instead must attend carefully to the intrapersonal range of variation as well as the range in the group. This does not deny that EI is empirically true. ". . . variability in achievement at the end of one year in the program is greater than the beginning placement variability; this seems particularly true for the early grades." (Robert Glaser, LR&DC Working Paper 2, 1966.) In context, Glaser's statement is compatible with an even stronger version of Principle EI, viz., that under school conditions the *more* effective the instruction, the *more* the increase in variability. But if brightness and dullness are not single traits, only one of which is possessed by every child, then we can work for a moral reconciliation between Principle EI and the claims of equality and justice in educational programs.

Hence the uncertainty about the status of Principle EI. It seems, from one point of view, to be rather a simple statistical fact: Anytime one increases the mean in a set of scores, the chances are that he will increase the range as well. From another point of view, EI seems one of the most solidly established empirical generalizations in the literature of educational research. From still a different point of view, EI seems a presupposition of some of our most valued practices in school life. From any point of view, however, EI is true. And its truth means that under normal conditions the individual teacher, facing a group of students who will learn at widely varying rates, must rely on *informal* means of applying standard devices, i.e., retarding some pupils, accelerating others, and enriching offerings just to preserve enough homogeneity to make effective instruction in groups possible. The task is ultimately hopeless, though some will manage to fail with much greater distinction than others.

If the NSSE yearbook is a straightforward account of what is known about individual differences, the ASCD yearbook is a hymnal praising the value of individuality. The volume was planned and written as a unit; there is a single theme which runs through the whole: The aim of instruction is the release of the full human potential of each individual child and youth:

> "Human potential is a multi-dimensional, many-faceted personal possession of the human spirit with which educators are concerned. It is the responsibility of school people to manage the educational content and the learning situation in such way that for each pupil these inchoate forces within him become awakened and released."[8]

---

[8] Robert F. De Haan and Ronald C. Doll, "Individualization and Human Potential," in 1964 ASCD Yearbook, p. 16.

I must remind you that the ASCD yearbook was a collaborative effort in which the committee as a whole took responsibility for all the chapters. There is certainly a difference in the degree to which this ideological commitment, as stated by the chairman of the committee, is central in various sections of the brief book, but it is never absent and certainly never contradicted. One can see how seriously this committee took their position by contrasting their view with another partly ASCD publication of about the same date. The yearbook committee, Doll et al., were willing to recognize that *whatever* had to be changed in order for "school people" to exercise their "responsibility" as defined, *ought* to be changed. But the ASCD was taking part in the Joint Project on the Individual and the School, along with the American Association of School Administrators and other groups which published *A Climate for Individuality* (Washington, 1965). The latter, unsigned statement makes it clear that schools are going to stay pretty much the way they are and

Such a view is potentially a revolutionary doctrine, though the yearbook committee did not look directly at the political implications of their ideas. With hindsight, we can see how radically different were the ideologies of the NSSE authors and the ASCD yearbook committee. The former seem to think about instruction as a matter of moving people along a particular track of learning. Individual differences are of great importance, for it is only when we understand the individual learner that we can know how far, how fast, and by what means we may move him down that track. On the other hand, the ASCD group was at least started on a line of thinking which would end up denying that there is any one track of learning down which all must go, affirming instead that there has to be a unique track for each learner, one set by that learner's unique potential. Thus when the ASCD committee refers to the principle of widening the range of variation, they do so in qualitative, not merely quantitative, terms:

> "Teaching is designed to make learners less like each other rather than more like each other. It is the teacher's responsibility to bring to learners his perception of what they can become, . . . in such a way that he releases the catch-spring of their desire to achieve the best that lies within them." [9]

---

that while a great deal more can and ought to be done for the individual, it ought to be done without seriously upsetting existing arrangements. "Schools are geared to working with children in groups. . . . Basically, in school each child must accomplish most of his personal development within the context of the group, which may promote individual development—or hinder it." (*Climate*, p. 29.) I am not prepared to claim either that the ASCD yearbook committee's view was unique or that it was typical of something else. I do claim that the somewhat hyperbolic-sounding language of the quotation in the text above was deliberate and distinctive.

[9] Ibid. p. 20. This is an awesome responsibility to lay on the backs of teachers. It is not surprising that in the actual school environment teachers find ways, even unconscious ways, of escaping it. Consider the beautifully documented *Studies of the Use of Programmed Instruction in the Classroom* (LR&DC Technical Report, no. 1). One of the questions that had to be studied was how teachers would handle the individual differences in students' responses to programmed materials. Richard O. Carlson puts it this way: "Some of these [teachers'] practices could be classified by the cynic as sabotage. The clearest example of this is that as a program progressed, the levels of individual achievement and the range in number of frames completed varied widely, and, as a means of 'correcting' this, several teachers were discovered either consciously or unconsciously to be pacing students. That is, teachers were actually restricting the output of the students proceeding at the

Thus two yearbooks entitled *Individualizing Instruction* contain a profound disagreement on what it means to individualize. Anyone with a dialectical turn of mind would suspect that such divergence is possible only if there is corresponding convergence on the meaning of instruction. Here is the strange part: I discovered not a single sentence in either book which either defines or describes instruction. There are words of praise for instruction; it is said to be the "heart of" and the "key to" good education; it is what schooling is all about. And so on. But what is it? How do we know it when we see it? I believe that the more remarkable difference between educational thought today and that of eight to ten years ago, when those yearbooks were planned, is that we cannot now take the concept of instruction for granted as they did then. We are happy to learn that children are more different from each other than we had suspected. But to us, that fact is not the most significant thing in understanding individualized instruction. The most significant thing is to understand the concept of instruction.

II

From the point of view of research, no two movements in education could have been more unlike than programmed instruction (PI) and the educational efforts of the Office of Economic Opportunity (OEO). The first appeal of PI is to the researcher. Most of the variables that are a problem for the design of research into teaching effectiveness can be eliminated by using only programmed materials in the activity of instruction. For example, if one wanted to study how and what individual differences affect receptivity to instruction, then one must hold instruction constant

---

fastest rates. The logic of restricting output of fast students is tidy and makes good sense from at least one viewpoint. Explaining the same troublesome point to five students who have encountered it at the same time is less time consuming than explaining the same troublesome point to the same five students as they encounter it at different points in time. For the teacher who complains that there is never enough time, this appears to be most efficient. In fact, insisting that all students move at the same rate, which is attempted in many classrooms, can be supported by the logic." (LR&DC Technical Report, no. 1, p. 126.) The lack of correspondence between what the ASCD yearbook committee took to be the moral obligation of teachers and what teachers actually do in the school context is itself an example of the committee's failure to consider the revolutionary *institutional* implications of the position they took.

while studying individual differences as independent variables and achievement as the dependent variable. Even better, one can test for interaction effects between personal variables and variables in the instructional materials, but only if it is known exactly what the instructional materials are, exactly how they are ordered in sequence, exactly how much redundancy they contain—in sum, only if the materials are programmed and packaged.

It is not difficult, then, to explain why the impact of PI was greater on the research activity in education than on the instructional program of the public schools. What may not be so obvious is why PI was so popular in the OEO. Let us recall the rhetoric of those days. We spoke of a "cycle of poverty" in which education played a very substantial part. Poor people received far less education than rich people, and their lack of education ensured their continued poverty. We knew that schools distributed their benefits in ways that consistently penalized poor children and (on a statistical basis) overrewarded middle-class children. We believed that there were several different sorts of factors operating to produce this phenomenon. One was pretty easily identifiable: The conscious or unconscious prejudices of teachers which made them respond more favorably and with higher expectations toward clean, well-dressed, well-mannered (and usually white) pupils and just the reverse toward "the other ranks." A wonderful thing about a programmed text is that it is color-blind and has no sense of smell. A programmed text is going to give a poor child exactly an even break, which was more than we could then believe true of 95 percent of all live teachers.

But there was an even more compelling argument for PI in the "war on poverty": If we could discover exactly what it was that a poor child lacked (or had) that a middle-class child had (or lacked) such that the poor child did poorly in school compared with his middle-class peer, then we could make sure that the poor got (or lost) whatever was necessary to assure equality in educational benefits. We recognized that expressions like "culturally deprived" or "disadvantaged" or "low IQ" were "cop outs"—they served as excuses for evading the responsibility of teaching the child. What one really needed to find out was why this particular pupil could not apply the basic rules of arithmetical computation to verbal problems, why this other pupil could not read a two-page story and then tell it in his own words, and so forth. In short, those principles of operant conditioning which lie at the base of all forms of PI force one to attend to the individual learner, his present repertoire of responses with the gaps all revealed, and to accept fully the responsibility to help him acquire that repertoire which is necessary to academic and vocational success.

Notice that operant conditioning, like any serious stance in psychology, is not only a technical theory to account for certain observable phenomena, it is also an ethical doctrine of some depth and nobility. B. F. Skinner has taken considerable pains to point out his view of the ethical and political significance of operant conditioning. For the "war on poverty," that psychological doctrine fitted with great comfort. It was radically egalitarian: all men, like all pigeons, are endowed by their Creator with the right and the ability to learn. It was radically individualistic: each man, like each pigeon, must learn at his own rate and with his own style. Just as Skinner and his associates took subtle differences in response curves (even from animals of a single carefully bred, laboratory strain) as having great scientific significance, so too did the educational theory of the OEO take subtle differences among children as significant, politically and pedagogically. From Head Start through the Job Corps and Upward Bound, the instructional programs of the OEO were consciously and deliberately individualistic as well as egalitarian.

For strategic purposes, we usually sought to conceal it, but our *feeling* was that the "war on poverty" must begin with the battle to abolish the public schools' monopoly on education. Not that we had any basic ideological battle against public education. Far from it. Had we been more conscious of the historical roots of our efforts, we could have taken our watchword from that great leader of the public school movement, E. L. Thorndike:

> "The study of the facts and laws applicable to all men by virtue of their common humanity gives education its fundamental rules for the control of changes in intellect and character. The study of the facts and laws of individual differences enables us to apply these principles economically in the case of each individual whom we seek to educate."[10]

---

[10] Edward L. Thorndike, *Individuality* (Boston: Houghton Mifflin Co., 1911), p. 3.

No, the conflict between the OEO and the public schools was not ideological but institutional. As it worked out in practice, we believed, the schools found it more economical to ignore our clients, the children of the poor, and to use their resources to advance just those individuals whom they really sought to educate, i.e., the children of the middle class. Had we appealed to Thorndike we would, in effect, have been recalling public education to the tasks set for it before it became a great system, a major industry which acquires its raw materials by coercion and is a protected market for its own products. Had we appealed to Thorndike, we would have linked operant conditioning to those psychological principles which antedated the time when low IQ scores were a respected alibi for failure to teach. But, like most New Frontiersmen, we were little concerned with ideology and history but greatly concerned with a power: So long as "they" (the public schools) had it, "our people" (the poor and their children) would suffer.

While I no longer believe that PI (or its offspring, Individually Prescribed Instruction, IPI) provides an adequate set of programs or principles for effective education, neither do I believe that movement to be ignoble or irrelevant. Surely it is not ignoble to affirm that everyone can learn. Nor is it irrelevant to insist that educational research focus on the individual learner to discover how he can be helped to learn better. You will remember that the funding policies of the OEO were frequently criticized for an "antiresearch" bias. There really was such a bias, of course; or, if you prefer, there was a strong determination that money allocated to improve the education of poor people not be used to further the careers of university professors and their research assistants. But from another perspective, most of the money spent on direct educational service to individual students had a research component built in. For if one *is* seeking to understand the "facts and laws" about *this* individual in order to help him learn more "economically," then one is engaged in educational research. Or, if you do not wish to call that sort of thing "research," it is still not an ignoble nor irrelevant thing to be engaged in.

Initially, then, the struggle between the OEO and the public school was to force the latter to give equal justice to the clients of the former. The schools serve as gatekeepers to all sorts of goods in this society—to vocational and academic skills, to further education and training, to jobs and income, to personally enriching culture, and so on. If every individual child is to have an equal chance at these goods, then the school must recognize each child for what he uniquely is. Like the NSSE yearbook, the initial thrust of the OEO was not to quarrel with the goals of public schooling, but only to insist that poverty not be allowed to prevent each individual child's having a fair shot at achieving those goals.

Later, of course, when the "war on poverty" had been lost and abandoned, the cry, "Schools are a middle-class institution," took on a more revolutionary meaning. Not only did schools not give kids a "fair shake" at the goods of this society; the goods that schools offered as rewards for conforming to its rules were shoddy, class-bound, pseudogoods not worth pursuing anyway. It was with a sense of despair that the politically disillusioned began to drift into the revolutionary doctrine which the ASCD yearbook had hinted at but never developed: To individualize instruction means to get out of the way of each pupil so that he may fulfill his own unique potential. At that point, of course, radical educational thought turned its back on PI and all other derivatives from operant conditioning. For all these ways of looking at education assumed that one could "specify terminal behaviors" (to use the jargon) toward which instruction is devoted. But if instruction is individualized in the ASCD sense, that is impossible: No one can determine in advance what any individual child must do to fulfill his own potential or what the end result will look like. From the political as well as from the psychological side, then, the old expression "individualized instruction" came to express not an ideal but a paradox. To the extent that an institutionalized encounter between a teacher and a student was an instance of instruction, it was false to call it individualized because the uniqueness of each child is ignored. To the extent it is individualized, it could not be instruction because terminal behaviors cannot be identified.

III

This is about where we stand today. There are really two competing ideologies centering upon the idea of individualization of instruction. One, which I will call the NSSE sense (recognizing, of course, that this is but a label and cannot be taken as descriptive of the attitude of all the contributors to the 1962 yearbook, much less all the members of the Society), takes the *problem* of individualization to be that of increasing the effectiveness of instruction by adapting it to the traits of the learner—his present repertoire of knowledge, the rate at which he can learn new material, and so on. The second, which I will call the ASCD sense (attaching the same disclaimers), takes the *problem* to be that of releasing the full human potential in each individual learner. Such different ideologies cannot be resolved by research in any conventional sense. Faced with such a situation, a philosopher is tempted to lend his support to one side against the other, but that would be a dereliction of duty. The philosopher's task is to find ways of posing clear questions that are *not* ideologically loaded toward one side or the other. I should like to come at this task by a preliminary analysis of the concept of *instruction*, a term which, strangely enough, is ignored in most discussions in this area. Then, in the following section, I shall try to show how the significant policy questions may be posed in nonideological terms.

A recent paper by Paul Dietl contains a most interesting suggestion—that there is a class of sentences in which it is impossible to substitute "instruct" for "teach" without rendering the sentence meaningless.[11] I wish to borrow that idea and to generalize it, perhaps beyond the point that Dietl would approve. I shall argue that there are two slightly different grammatical forms in which we talk about instruction, and that only one of these forms can be expressed in the standard paradigm of "teach." If these propositions can be supported, they provide the basis for a somewhat clearer notion of what we mean by "instruction." And if the concept of instruction becomes clarified, then the problems for empirical researchers might also become clarified.

Let us begin, then, with the standard paradigm of "teach."[12]

---

[11] This analysis of the concept of instruction is borrowed from Paul Dietl's unpublished paper "Teaching, Learning, and Knowing," a paper containing such a wealth of new ideas that its author can afford to lend a few to his friends.

[12] There is a wrinkle in the paradigm: When "X" is the *name* of what is being taught (rather than what is being taught), then the paradigm may read:

A teaches X to B

A	teach[es]	B	X
	(tensed verb)	that	
		to	
		how to	
		when to	
		where to	
		whether to	
		about	
		whatever	
		(etc.)	

The conceptual problem, of course, is to show exactly where the limits are of the values which may be given to A, B, and X such that the paradigm produces a meaningful sentence.[13]

The customary approach to a problem of this sort is to start right in the middle, i.e., with those values for A, B, and X which unmistakably yield a meaningful sentence. The custom has much to commend it. If A and B name (or provide a role description of) two persons possessed of all the normal faculties of human beings, and if X is either an intelligible proposition (following "that") or a verb describing a voluntary action (following "to"), clearly the paradigm produces a meaningful sentence. Equally clearly, one can produce a formula to guarantee that the paradigm generates gibberish: Let A and B be trigonometric functions and X name a color; there is no purely syntactical move which can save the resulting sentence from meaninglessness.

The term "instruct" is a more complex term than "teach" and "train" and "indoctrinate." In one sense of the term, of course, "instruct" is only distantly akin to "teach" and much closer to "order" or "direct." That is to say, there is nothing very pedagogical-sounding about a report like, "I instructed my butler to announce my absence to all callers." Whether one says "instructed"

Thus we *may* say: "George taught the pythagorean theorem to his fifth-grade class." But we *must* say: "George taught his fifth-grade class that the square of the hypotenuse of a right triangle. . . ." The purely grammatical distinction between dative and accusative cases is useless as a clue to the logic of the paradigm.

[13] There is a further problem to show what are the necessary and sufficient conditions for that sentence to be true, but we cannot touch that further problem here.

or "ordered" seems to depend on whether one spent his time in the parlor or on the parade ground.

But even in a context most rank with coercion, there is a remnant of rationality necessarily present in "instruct" which may be absent in some other pedagogical predicates. One can train or teach (here synonymous) a parrot to utter the words "He's not at home" whenever the doorbell sounds, but one cannot say, "I instructed my parrot to announce my absence to all callers."

We have been picking a tortuous path through the trivial to establish one significant point. In the nonpedagogical sense of "instruct," if A instructs B to X, then B has a *reason* (albeit not by this fact alone a sufficient reason) to X. B must then be an entity capable of understanding, seeing, grasping the point or purpose in *his* (B's) Xing, even if that point or purpose is only to incur A's favor or avoid A's wrath. In that minimum case, A must have authority if he is to instruct B; i.e., A must have some power such that B objectively has an interest in incurring A's favor and avoiding A's wrath.

There is a hint of support for my view here if we consider how it is that instruction in this nonpedagogical sense may fail. When teaching fails, we say it was unsuccessful. When an order fails, we say it was disobeyed. When an instruction fails, we say it was disregarded. Only a creature capable of self-regard could fail, in this way, to carry out an instruction.

This minimum sense of "A instructs B to X" has given a bad name to the whole concept of instruction and may explain why the yearbooks ignored the concept altogether. No matter how many values one gives to X, even if the instruction never be disregarded, one could never account for B's becoming an educated man by referring to a succession of "A instructs B to X." One can account for efficient servants or effective military operations by referring to the quality and quantity of instruction (in this nonpedagogical sense), but a liberal education is something else altogether. Notice: No matter how individually instructions are given to the butler, they only make him a more perfect butler, not an educated man. And there is also the danger that as the effectiveness of the instruction (still in the nonpedagogical sense) increases, so also increases the likelihood that B will become conditioned to depend on A as an external source of standards and values. And so on. It is hard to imagine that we could eliminate nonpedagogical instruction from all social life, but surely we may think it has no necessary connection with education. But then, what place does *any* instruction have in education? Responsible educators have been very hesitant to provide a definite answer.

As we now treat of the pedagogical sense of instruct, we must keep in mind the central canon of its nonpedagogical use: When A instructs B to X, B can understand some (self-regarding) point or purpose in Xing. Usually when we talk of instruction in its pedagogical sense, we say A instructs B *in* X, using the same locution found in training and indoctrination. In addition to instructing my butler to announce my absence to all callers, I can also instruct him in how to do it: Like *this* to apparent creditors or CIA agents, like *this* to old cronies, like *this* to familiar young ladies, and so on. If I am instructing him in how to do it, then *I* must believe that there is point or purpose in doing it this way. *That claim is purely conceptual;* it would make no sense to affirm that I am instructing him in (here expressible as teaching him) how to do it and then denying that there is *any* point or purpose in doing it this way. The next step in the argument, however, is not conceptual but pedagogical: To the extent that he understands the (nonself-regarding) point or purpose of doing it this way, he will be more able to learn what it is I am instructing him in how to do.

This pedagogical point may in many instances be false in fact. It may be that A can instruct B in how to X more efficiently by *not* revealing the point or purpose of Xing. But as a pedagogical point, the premise is not merely a factual claim, but also and more importantly a moral claim. When we say that B will learn better if he understands the point or purpose of his instruction, we might mean (in many instances we ought to mean) not that it is more efficient, but that it is morally better if he learns by and while understanding the point or purpose of his instruction than if he learns in some other way. Remember that it is neither a contingent fact nor a moral claim but a conceptual necessity that there *be* a point or purpose to his instruction.

To recapitulate before moving to individualizing instruction: There is a sense of "A instructs B to X" which is nonpedagogical, i.e., can never be expressed in the paradigm of "teach." There is another usage of "instruct," i.e., when we say A instructs B *in* X[*ing*], which is compatible with the paradigm of "teach"; this second usage of "instruct" (carrying over part of the first sense) inhabits a very peculiar region of the broad expanse covered by "teach," i.e., the region where if it is true that A teaches B *in* X, then it is also true that A, at least, believes that there is point or purpose in X. Although with increasing attenuation as we move to more abstract matters, whenever we can use "giving instruction in X," for "teaching X," we imply (with either expression) that we believe there is a point to X[ing]; we put ourselves under some obligation to say what that point or purpose is, if asked.

Exactly where that region lies in the land of "teach" cannot be stated in general; one has to test in each particular case. "A instructed B in the basic chronology of American history" sounds a bit pretentious as a way of talking, but it is certainly compatible with "A taught B that Columbus's first voyage to the Caribbean preceded Vespucci's." Assuming that A is enough of an authority to give instruction in chronology, then the "A instructed B" statement is not compatible with "A taught B that Vespucci's voyage preceded that of Columbus," though the latter could well be true and describe the intentional action of an expert in chronology. But doing so would violate the point or purpose of *instructing* in (or teaching) chronology.

Or consider another case. You are giving instruction in how to translate a story read in a book into the reader's own language. Suppose you teach some youngsters to write on a piece of paper a list of questions, beginning with "*Who* was the principal character?" "*What* did he (she) do?" and so on. Suppose you taught the youngsters to answer these questions after reading the story. Suppose that neither you nor the students see any point or purpose in these exercises. How then could you use the verb "instruct" (in its pedagogical sense) to describe that activity? I think it quite clear that we could say that you had taught them something; we could say without embarrassment what it was that you had taught them. But would we say that you had given them instruction in (or taught them) how to translate stories into their own words? At that point we should be quite chary. That *sort* of test, I believe, can be applied whenever one wishes to see whether he is treating of that region of "teach" where one can equally well use the term "instruct."

This way of defining "instruction" as a rather particularized region within the broad concept of "teach" is, I admit, a rough beginning on a concept that needs a great deal more attention and research. Responsible educators from kindergarten teachers through college presidents insist that instruction, typically "a sound instructional program," is the central core of the educational endeavor. It is here that teaching must necessarily have point or purpose. It is here that the instructor's authority *over* his pupils is grounded in his *being* an authority in what he teaches. Instruction is what gives point and purpose to any school, as educators have long claimed without being entirely clear what they meant by instruction.

What is *not* implied by "instruct" is equally important. "A really competent instructor" is about as superlative as one can get in using the term to describe a person. But we do talk about truly great teachers, implying attributes of personality that go far beyond competence as an instructor, though the latter is probably a necessary condition for being a great teacher. In many parts of the world, students' protests have listed increased competence in their instruction among their major demands. In this country, "closer faculty-student relations" is more likely to be heard than "competent instruction." The very neutrality of "instruction" explains in part why, in the ideological struggles over educational change, the concept has received so little attention. But if the argument in this section is correct, it does give a window through which we may see that ideological struggle more clearly.

IV

Let us summarize and recapitulate: Section I pointed out two different things people might have in mind when they speak of *individualizing* instruction. The first, or NSSE sense, focuses on the problem of preventing variance (among the individual members of an instructional group) from interfering with the effectiveness of instruction. The second, or ASCD sense, is concerned with helping each individual to release his unique potential as a human being. In the first sense of "individualize" we know pretty well what steps we would have to take if we were serious about individualizing instruction, even though we may be doubtful that we could secure the money, the talent, the techniques, and (more importantly) the political power necessary to take those steps. But we are not clear, and the ASCD yearbook really did not help us become clear on just what instruction contributes to individualization in the second sense. We are, or should be, willing to grant that a society which exalted individualization in the ASCD sense would be a noble society indeed, one worth striving for. But we are not exactly sure how or whether instruction would be practiced in that society.

So we had to look more closely at the concept of instruction itself. We found it to occupy an important region in the great domain of teaching. But this region is shunned nowadays by fashionable travelers seeking educational reform, for therein warm human relationships and existential encounters are found only by accident; what is necessary to instruction is, rather, rational purpose and authority. Particularly those committed to the ASCD sense of individualization are unlikely to find anything very worthwhile in instruction, even though they continue to combine the terms in their talk.

Now when we ask about the connection between "individualized" and "instruction," we find that the answer is not to be found solely in casual alliteration nor conceptual interpenetration. In addition to being a distinctive pedagogical predicate, instruction belongs in another category of actions which persons address to other persons, actions like helping, rewarding, encouraging or punishing, threatening, torturing, insulting, and others. These are all actions which can be done more effectively if they are individualized in the NSSE sense. Thus if Jones is a member of a class which I punish as a whole by denying its members access to the playground, my action may be no more punishment for Jones than the briarpatch was for Brer Rabbit. If I am to punish Jones, I must know what sort of thing Jones *finds* hurtful. I could injure him, even grievously, without his being aware of it, but we would not call that punishment.

To torture effectively, I must know what causes pain to my prospective victim. There is sufficient similarity in the neurophysiology of human beings that I can apply general methods with reasonable assurance that what causes pain in general will cause pain to any particular victim. Or, I might have a perverted desire to design exquisitely individualized torture which causes extreme pain to Jones, although most persons would find the experience pleasant or at least neutral. At the other extreme, I can reward Jones only if I know what sorts of things *he* finds rewarding. Again, general guides are usually trustworthy, but care and thought and knowing Jones can ordinarily make the action of rewarding Jones more effective.

For an action like encouraging Jones, the applicability of general guides is more restricted. "Come on, Jones! You can make it," may be encouraging; but again, surely, it may not be that at all. If Jones has found that that sort of prod has usually been followed by his failing to make it and being punished for his failure, then the hearty "encouragement" may well be just the opposite for Jones. Encouraging is the sort of thing which we individualize not just to make it more effective, but rather to make sure that we are doing *that* and not the reverse.

Now what sorts of things would we have to know about Jones if we are to individualize any of these actions toward him? The obvious fact is that we would have to know something specific about him for each of these specific actions. To punish—what he finds hurtful. To torture—what causes him most pain. To reward—what provokes feelings of accomplishment (and, possibly, gratitude). To encourage—what recollections or prospects most conduce to feelings of confidence and willingness to risk all-out effort. To instruct—what he is ready to learn, meaning both what he now knows and what next steps he is willing to take. Notice that knowing these things *about* Jones is not a sufficient knowledge-condition for individualizing the action toward him. The agent would also have to have the necessary technical skills and apparatus to know how to individualize any one of the actions mentioned.

Now one can look back at the various kinds of research mentioned in Section I of this essay and see that they classify very easily by the kinds of problems that arise when we consider the two sorts of things we have to know if we are to increase the effectiveness of instruction by individualization. We have to know some things about the person being instructed. If he is six years old, then, in general, what are six-year-olds ready to learn? What steps can we take to learn more about this particular child? We also have to know how to use the solutions we get to our first problems to design instructional techniques, either for particular learners or for classes of learners who are sufficiently similar in relevant characteristics that a single instructional technique can be applied effectively to a class. In practice, our way of solving both these problems may appear quite primitive compared with our solutions to formally similar problems in, say, industrial production. But the principles are fairly clear, and there is little reason to doubt that practices will become more effective as research proceeds.

But another whole order of questions is raised if we take the ASCD concept of individualization seriously. Look again at those actions which are like instructing in that they can all be individualized in the NSSE sense. About all of them one may ask: By what right do you intervene in another person's life by Xing him? If we fill in the blank with punish, torture, insult, and so on, the question gets its foothold from the fact that these are in themselves malevolent actions. Anyone who engages in them must recognize that such actions are rightly subject to moral condemnation unless some clearly overriding condition can be offered to justify them. The moral sensibilities of most of us today are such that we have great difficulty imagining any condition which would give a person the right to torture another. We vary widely in our beliefs about the conditions which would give a person (acting individually or for a collectivity) the right to punish another. And so on. These malevolent acts are very different in many respects, but they all share these two features: They can be increased in effectiveness by individualizing them. And we recognize the appropriateness of the question: By what right do you engage in that action?

But is malevolence a necessary condition for an action's being subject to the question: By what right do you do it? The obvious answer is no. I observe Miss Jones trying to make her way through an obstacle course. I shout: "Come on, Jones! You can make it." A bystander doubts my benevolence; he questions: "By what right do you mock that poor girl who's making such a valiant effort?" The bystander's question is effectively stopped when I convince him that I really was encouraging Miss Jones and not mocking her. We call it encouraging because (among other things) it is a benevolent act toward Miss Jones.

Suppose Miss Jones does make it and *she* pants up to demand: "By what right did you encourage me? I was trying to find out whether I could make it without encouragement from you males." Miss Jones's question is not stopped by, "But I was encouraging you."

There would seem to be no rules for this sort of thing. One's general duty to help, encourage, and otherwise act benevolently toward other persons is always implicitly qualified by such phrases as "rather than the reverse" *and* "when it is fitting or appropriate to do so."

The question of whether one has the right to intervene in another person's life by instructing him in X is not answered by saying that, in general, instructing is a benevolent way of acting, rather more like encouraging or rewarding than like punishing or insulting. Let us grant that it is, in general, benevolent to instruct. (The point could be debated, but that is another issue.) But surely one is not thereby given a right to go around instructing everyone he meets on anything one happens to know about. It is bad manners to be, as we say, didactic in conversation. What is it when adults set up a compulsory school system which requires children to receive instruction for hours at a time?

This is, I believe, the sort of question which is raised whenever one thinks about the ASCD concept of individualization. Indeed, its *statement* of that concept fits most tellingly with just the question we are now considering. Thus Q: By what right do you intervene in the life of that child by giving him this instruction? (Let us assume that the manner of instruction is benevolent and its matter not morally wrong.)

A_1: In order to release his full potential as a human being.

This is not to say that A_1 is the only response which can answer Q. We can imagine a very large number of answers, each of which is more or less persuasive.

A_2: So that he will be able to get a job when he grows up.

A_3: To give him the background needed for the next step in his intellectual development.

A_n: Because if we do not so instruct all the children in this society, we will revert to a band of murderous apes—and he's one of the children.

But our intuitions warn us that A_1, the ASCD answer, is in some important respects different from the other answers, each of which is, by the way, a relevant response even when not conclusive. Just how is A_1 distinctive? In two closely related respects: (a) In contrast to other possible answers to Q, if we could establish the truth of the implicit empirical claim in A_1, it would be a conclusive, unpursuable answer to the question of right. The implicit claim is: The instruction I am giving to this child is necessary to the release of his full potential as a human being. Anyone who should grant that claim and still say: "Yes, but I don't see how that gives you any right so to instruct him," would show that he does not know what it is to ask and answer questions like Q. But (b), we face enormous difficulty in establishing the empirical claim in (a). What would we have to know? How, in God's name, could we ever gain adequate evidence that this particular matter in which we are instructing him is *necessary* to the full release of his potential? We do believe that some things are necessary to the release of human potential, such as a child must learn to talk, to move his limbs at will, to recognize other persons, and so on, if his *human* potential is to be actualized. But those are not the things in which we give *instruction*. Could we even conceive a research design which could establish for any particular child that *this* instruction is necessary to release *his* potential? If there are such research designs, in any conventional sense of the term, then I have not encountered them in my unsystematic perusal of the literature on individualized instruction.

If the argument so far is correct, it brings us to this conclusion: The distinction between the NSSE and the ASCD concepts of "individualization" is basically that each answers a different question. The NSSE concept and all the associated research is an answer, more or less convincing, to the question: How can we adapt instruction to the unique traits of each child so as to make instruction more effective? The ASCD concept, for which we could not even imagine supporting research, is to answer Q: By what right do we give instruction? In short, the NSSE approach focuses on the *how* question, while the ASCD approach, pushed far enough, forces us to consider *whether or not* we ought to engage in the activity at all.

This distinction, clear enough thus far, still does not deal specifically with instruction. We could ask the same sorts of questions about any benevolent intervention in the lives of others. But instruction, please recall, is that sort of intervention in another's life which provides him some purpose in acting a certain way or some reason for believing something he might not otherwise have believed. There is no obvious conceptual connection between instruction and individualization in the ASCD sense; indeed, as mentioned earlier, they may appear conceptually antithetical. For surely, if the point the child sees in my instructing him in Xing is that his Xing will secure my favor and avoid my wrath, that instruction has not advanced his unique potential in any significant way.

But if any person, in pursuing his own purposes, discovers that these purposes cannot be accomplished without his learning (that, to, how to . . . and so forth) something, and if he then seeks instruction in that something, the person giving that instruction has a plausible answer to the Q question. Let us call that answer A_p, the subscript in recognition to its similarity to the pragmatic philosophy of Dewey et al. A_p: Because he could not achieve his purposes without instruction, and he asked me to provide it.

A_p is a considerable advance over "Just because he asked me." For I must acknowledge a defect in responsibility if I waste a child's time by instructing him in idly entertaining matters which distract him from his purposes. And A_p is safe from that objection. But A_p is still not as compelling as is A_1. If the seeker of instruction has genocide as his purpose, I do not have the right to instruct him in the techniques of mass poisoning. But A_1, notice, is safe from that sort of attack. For the idea of "human potential" has built into it some distinction between humane and inhumane purposes. There is, of course, no guarantee that the norms of any given social order will be supported by instruction justified by an A_1 answer: But if there is conflict, the norms of the social order (and not the instruction) are thereby shown to be morally suspect.

If we now eliminate immoral purposes, what do we have to add to A_p to make it as compelling as A_1? May I suggest a premise

(P_p) which is itself to be justified in philosophical anthropology or general theory of human nature. P_p: A necessary condition for the release of any person's full potential is that that person have the knowledge and skill to pursue his own purposes.

Doubtless there are more elegant and eloquent formulations of this idea than P_p. But it will do for this argument. Its truth, as I said, would have to be established in a theory of human potentiality, but let us just suppose that P_p is true. If so, its utility is enormous. For A_p and P_p together *imply* A_1. Thus we do not have to look for the impossible to conceive research to validate that our instruction is a necessary condition for the release of this particular child's human potential. If we are reasonably careful, we can find out what purposes he holds independently of our instruction. We can help him discover what knowledge and skills beyond what he now has are necessary to achieve those purposes. We can help him decide whether or not he still wants to pursue those purposes now that he knows what he will have to learn in order to carry them through. If he then says, "O.K. Teach me how to read," the right to instruct has been totally vindicated.

It is here that instruction differs from other ways in which persons intervene benevolently in the lives of others. For instruction has a point or purpose independent of the purposes of those engaged in it. And persons have purposes which are independent of whatever instructional activity, if any, they happen to be engaged in. But *sometimes* it happens that the purpose of the instruction, some gain in knowledge or skill, supplies a necessary condition for an individual's achieving *his* purposes. And sometimes it happens that the person undergoing instruction has adequate grounds for believing that what he is learning is necessary to the achievement of purposes which he holds independently. In that case we could justify our instruction by appeal to the ASCD sense of individualization, i.e., supplying a necessary condition for the release of this individual's full human potential.

Now would we want to call that activity individualized instruction? It might be confusing. For the ASCD moral justification might apply equally as well to a lecture course taught to a thousand students at once as to a tutorial in the teacher's office. If the purpose of the instructing is made clear, if the individual deliberately engages as a student in the instruction because he believes that instruction is necessary to his purposes (which are also clear, it must be assumed), then it really does not matter *how* the instruction is given. A television set will satisfy *that* demand just as well as a warm, loving, humane teacher—though the latter may be able to supply other demands which a television set cannot.

I am loath to recommend that we change our language and call instruction justified on this ASCD criterion *individualized* instruction; it might sound strange. But we ought to be clear on just what is implied in any particular claim that instruction is individualized. In the NSSE sense, there is no more moral justification for individualized instruction than for any other sort of gratuitous intervention in the lives of others. This is not to say that instruction is not justifiable, only that it is no justification to say that it is individualized. In the ASCD sense, one ought not claim that instruction is technically effective just because it is morally justified by appeal to students' purposes. It is a question to which research could make an important contribution by asking whether congruence with students' independent purposes makes instruction more effective. One imagines that there are conditions under which student purposes do constitute an important variable in determining effectiveness of instruction and other conditions in which they are of little significance.

But when we look at the truly monumental questions of educational policy that must be answered in the coming decades, our first premise ought to be a frank recognition of the moral inadequacy of the NSSE approach, taken by itself. That conclusion seems to me the single important outcome of the present inquiry. We simply must not be led into the false belief that instruction which is individualized for technical efficiency is ipso facto a morally justified intervention into the lives of students. That false belief is easy to fall into. We tend to feel that the moral criterion posed by the ASCD yearbook has been satisfied if the size of the instructional group has been reduced to one. Surely young Jones's unique purposes and potentialities *are* the determinants of his education if he is receiving one-to-one instruction—even if his instructor is a computer. But in practice, the exact opposite is more likely true. The logistics of individualizing instruction are very difficult to handle.

For the foreseeable future, the more effectively we individualize in the NSSE sense, the more logistical constraints will determine the matter and manner of instruction; thus, the less room there will be for an A_1 answer to the question of our right to instruct at all.

So we face policy issues with a number of variables which are empirically dubious and morally ambiguous. One wishes we could return to Thorndike's confidence in the power of a scientific method to give us both the ends and means of education. But we can't.

CHAPTER VIII

Conceptions of Culture and Society in Educational Research on Individuals

BRIAN HOLMES

Several approaches informed comparative education before World War II. The first was developed by historians who wished not only to describe and compare educational systems, but also to understand, using historical modes of explanation, why in different countries educational policies and practices are as they are. Another approach involved the collection year by year of comparable statistical data and policy statements so that trends of development could be discerned and compared. A third approach, which emerged in the late nineteenth century, made explicit the view that education could not be successfully studied in isolation but only in its relationships with other aspects of society. Of course, those who hold that sociology is an extension of history would maintain that the historical and sociological traditions in comparative education are almost the same.

Interest in comparative studies in education using sociological techniques gave rise in the 1950s to cross-national studies of relationships between educational opportunity and social-class background. Somewhat later, cross-national research attempted to show how societies and individuals benefit from an investment in education. Finally, while in the past rather little attention was paid in research by comparative educationists to individual attitudes and achievements, in the last decade major researches using psychometric techniques have studied individuals in different national systems of education. Thus, since the war, sociology, economics, and psychology have contributed to comparative education.

One value of these cross-national studies is that they reveal very clearly some of the difficulties inherent in the scientific study of society and individuals, in particular by showing the extent to which concepts of culture and society are implicit assumptions in educational research. Frequently these are taken for granted in national studies; they cannot with safety be ignored in comparative research.

It is evident, for example, that valid tests of attitudes or achievement cannot be constructed without reference to conceptions of culture and society, that is, if they are going to measure what they are intended to measure. An emphasis on developing reliable tests in cross-national research may be misleading, since reliability does not necessarily mean that the tests are valid. Claims that experimental or empirical research in education increase validity and reliability are not infrequent, but confidence in such investigations has ebbed and flowed, perhaps because on the one hand nonexperimental wisdom has apparently failed to find workable solutions to pressing problems. On the other hand, at times there has been an inordinate dependence on empirical inquiries in attempts to increase the rate and amount of progress in understanding individuals in society.

So in the 1920s experimental techniques drawn from psychology became the vogue. Sociologists then came along to step up the amount of research on individuals and social institutions. The results of these investigations have increasingly been used to formulate educational policy or as arguments in favor of accepting this or that policy—sometimes with disastrous consequences.

In contrast, historical and philosophical modes of inquiry in education have tended to fall into disrepute. The kind of criticism made of the history of education by Lawrence A. Cremin and Bernard Bailyn may account for its lowered status. As for philosophy, the application of analytical techniques clarifying basic educational notions and arguments gave a new direction to research in this field of education. However, that "pure analysis" may fail to inform vital decisions explains the limited influence such philosophers exercise on policy in contrast to the ideologically committed dogmatists. In addition, the tendency of some philosophers to overlook the need to subject conclusions drawn from philosophical constructs to the test of experience accounts for criticisms of the relevance of philosophy in educational research.

Comparative studies show that the dichotomy between empirical (experimental and nonexperimental) and nonempirical (qualitative) research is false. If research on individuals is to be useful, both elements are needed. In this paper some of the claims made for empirical research will be examined. The first of these is that it is objective and, by implication, that qualitative analysis is not. Another claim is that empirical studies alone can provide knowledge about nature, society, and individuals, all else being a matter of opinion. A further claim is that its results are more reliable than those derived from philosophical inquiry. Against these claims it will be argued that conceptions of culture and society should be analyzed philosophically in order to improve the validity (and particularly the construct validity) of experimental and nonexperimental empirical research. It will also be argued that philosophical inquiry which does not submit itself to some of the tests of practical life is bound to look arid and irrelevant.

Some Debates in the Philosophy of Science

As a basis for discussions, dictionary definitions may give rise to verbalism. Nevertheless, the current meaning of words may well be judged against more traditional definitions. For example, according to the *Shorter Oxford English Dictionary*, the word *empirical* has broadly two meanings. Both are drawn from the mutual practices of a sect of physicians who based their rules only on experiences. One derived meaning of empiricism relates to practices ignorant of scientific knowledge. According to this definition, the empiricist is a quack or a charlatan. The other derived philosophical meaning relates to the theory which regards empirical data as the only source of knowledge. The *Shorter Oxford Dictionary* illustrates the philosophical context of the word by referring to John Stuart Mill. And *Chambers Technical Dictionary* clearly identifies "empirical" and "empiricism" with those theories of knowledge and scientific method known as induction. Hence: "Empirical: Said of a rule or generalisation which is induced solely from observation, without correlation with other scientific laws from which critical phenomena are then deduced for experimental observation." A nominalistic definition would associate empiricism, at least traditionally, with theories of induction propounded by Aristotelians Bacon, Locke, and J. S. Mill and subscribed to by creative scientists such as Darwin, but criticized, among others, by Berkeley and Hume.

In practice the implication is that empirical procedures should be concerned with "facts" and "observations" based on experience, and that "facts" can be observed objectively only if they are quantifiable—as though the instruments for gathering and measuring such data are themselves simply derived from experience. Both the explicit and implicit meanings ascribed to empiricism have been criticized by philosophers of science, so that claims by empiricists that their research is scientific should be viewed with caution.

"Empirical" as meaning "induction" has been attacked by philosophers of science and practicing scientists. K. R. Popper's analysis of it and his categorization of Mill as one of the many nineteenth-century historicists are well known. Popper's suggestion that the hypothetico-deductive method (put forward by Mill in Book VI of *System of Logic* when he found that he could not apply the methods of the natural sciences as he saw them to the social sciences) more accurately describes the processes of scientific inquiry has commanded wide support. A distinguished practicing scientist, P. B. Medawar, after criticizing some of the assumptions of induction, concludes by writing, "the hypothetico-deductive system seems to me to give a reasonably life-like picture of scientific enquiry, considered as a form of human behaviour."

Medawar acknowledges his debt to Popper as the principal modern advocate of the hypothetico-deductive theory. Its main constituents are the nonlogical formulation of some hypothesis which is then exposed to criticism, usually through experimentation, which makes use of logic and empirical testing by comparing observations and experience with the logical consequences of belief. If events logically predicted from statements occur, confidence in the hypothesis is strengthened; otherwise it is abandoned altogether. The empirical element in science occurs when hypotheses are confronted with the test of experience and subjected to comparative evaluation.

One significant difference between induction and the hypothetico-deductive methods of inquiry is that while in the former hypotheses are induced only after the careful collection and classification of observable data, i.e., "facts", in the latter the source of the hypothesis is often intuitive. Today the controlling influence of a hypothesis in the collection and classification of data, i.e., colligation, is so widely accepted that it need not be debated here. While in Mill's theory an inductively reached tentative hypothesis (of limited validity and application) can through successive stages reach the status of a law unconditionally valid in time and space, in hypothetico-deductive theory, laws, like hypotheses, always remain hypothetical, to be confirmed or refuted in a specific context.

As for the objectivity of "facts" and testing procedures, one issue turns on how far observations can be construed as independent of conceptualization. The view that "facts" or sense data are, or can be, received by the mind without the latter imposing interpretations or constructions on them has been criticized. On the other hand, the complete absorption of "facts" into theoretical constructs is not acceptable. P. K. Feyerabend, for example, argues in favor of the relative autonomy of facts by saying that they exist and are available regardless of theories to be tested, but that the description of facts depends upon some theory. He goes on to say that some facts can be brought to our attention only with the help of alternatives to the theory to be tested. In like vein, Israel Scheffler draws a distinction between the independent existence of facts and theories and the isolation of one from the other. In actual thought there are indissoluble links between fact and theory, but for theoretical purposes analytical distinctions can be made between them. Such arguments lead to the conclusion that empirical evidence does not consist of "facts," pure and simple. There is a sense in which data or facts are analyzed, manipulated, or, indeed, manufactured in the light of some theory.

Operationalism offered an earlier solution to this problem by maintaining that theoretical concepts should be defined in terms of measuring operations, which are specific and unambiguous. In view of the confusion created in physics when verbal definitions of mass, length, and time were questioned at the turn of the last century, it is not surprising that operationalism grew out of the work of a physicist, P. W. Bridgman. In the social sciences similar problems of definition have made operationalism an attractive solution, particularly in psychology. In general, however, willingness to accept only quantified data in the social sciences is a trend towards operationalism. It is a tempting but naive solution—tempting because, provided co-workers accept the defining operations, reliable testing instruments can certainly be developed; naive because the test may not be valid.

Thus, IQ tests adequately define intelligence only if conceptions of intelligence are not in dispute. In the same way, essay-type examinations may perfectly well define achievement provided there is agreement on what constitutes achievement. In brief, operationalism can improve reliability without touching the problems of validity in educational research on individuals. This becomes very apparent when operational definitions are employed in cross-national research.

Examples of such research referred to later illustrate in a practical way some of the points made by such critics of operationalism as Carl Hempel. Take the measurement of "achievement." This is a culture-centered term which if used scientifically requires that its systematic function within a set of theoretical principles and terms should be known. At the moment there is no internationally agreed upon set of principles which make it possible to give unambiguous meaning to "achievement." A second objection to operationalism is that no scientific term can be synonymous with one set of operations because the latter afford criteria of application only within a limited range of conditions. National conditions under which "achievement" is measured differ greatly. Thus, tests of educational achievement depend for their validity on national sets of theoretical principles and on the national conditions in which they are applied.

This is to accept Popper's conclusion that measurement presupposes theories and, as he says, no operation can be adequately described in nontheoretical terms. IQ tests presuppose some theoretical concept of intelligence; achievement tests are derived from theories of "what ought to be known." Measures of educational productivity depend upon concepts of what constitutes "education" and theories of productivity. Crude operationalism, if it denies the contribution of prior and culturally determined theories, is an unsatisfactory basis on which to build research in education. It is even less satisfactory as a basis for cross-national investigations.

On the other hand, if the revolution in scientific thinking culminating in Einstein's theories exploded belief in scientific certainty and objectivity, the dangers of denying the relative autonomy of facts should not be overlooked. While the extreme view of some philosophers, such as Sir Arthur Eddington who rejected a belief in an objective real world in favor of conventionalism, commands little support, the objectivist-subjectivist debate as continued within the Vienna Circle during the 1930s still has its influence. Scheffler discusses it at some length in *Science and Subjectivity*. Otto Neurath's theory that the idea of checking beliefs (or predictions from them) against experience is misguided lends support to philosophers of education who accept the coherence and self-consistency of a system of statements as the only tests of its value. This position is as untenable as that of some present-day empiricists who maintain that all knowledge stems from experience. Nevertheless, several features of Neurath's point of view are extremely relevant in cross-national research. For example, observations of facts are only possible through a shared observational language. Beliefs are frequently checked against purported statements of experience—that is to say, not against observation, but against observation reports. These warnings are particularly relevant when interview and questionnaire techniques are employed in educational research.

The thesis advanced in this chapter is, then, that in educational research on individuals the beliefs and theories which inform it should be examined critically for coherence and self-consistency. Clusters of conceptions of culture and society are among those theories which should be so examined. Only after such analysis is it possible to establish reasonably valid measuring instruments. If these are then used to test hypotheses, the purpose of testing, i.e., confirmation or refutation, should be clear. And investigators should be aware that testing procedures may involve checking beliefs against observation reports rather than against experience. Finally, acceptance of the hypothetico-deductive theory of scientific method implies that some problem stimulates research workers. It should be analyzed; then, prior theories (or hypotheses) direct attention to particular kinds of data, enable facts to be analyzed, manipulated, or manufactured, and precede the establishment of measuring operations. If these several intellectual tasks are not accomplished before nonexperimental empirical tests are administered, the validity of the latter must be seriously questioned.

Theoretical Constructs in Educational Research

If we accept that educational measurement presupposes theories—that theories and hypotheses are intimately involved in empirical inquiries—then it is necessary to ask what set of theoretical principles is most often assumed in educational research. How can these principles be identified, made explicit, and classified so that they can be used in educational research? Is it possible to establish sets of theories which are useful in cross-national and comparative research? How can systems of coherent statements of principle be established? In this section an attempt will be made to answer these questions.

First, however, it should be noted that educational measurements are made in a social and cultural context and that all the artifacts of such a context (or environment) should be known. These may be usefully classified as follows: First, there are normative statements, i.e., assertions of what ought to be the case. These may be formulated as theories or hypotheses or they may be regarded by some philosophers as corresponding with reality, i.e., a priori true and untestable assertions. Into the second group of artifacts may be placed social institutions of which the school itself and, within it, the curriculum are examples. Not all sociologists agree that the distinction between norms and institutions is sound, but it serves my analysis well. A third feature of the environment includes geographical conditions such as terrain, climate, proximity to the sea, latitude, longitude, and other physical conditions over which man has little or no direct control. Such a classificatory system cannot include everything, although theoretically for research purposes this is desirable. Selection has to be made.

This paper is concerned only with what I have called elsewhere "normative patterns." These are the theoretical principles which inform empirical research in any country; they are made up of normative statements, i.e., assertions of what ought to be the case. However, they may be held to describe "what is" rather than "what ought to be." Popper's view is accepted here that they are man-made statements which one is free, once one recognizes them, to accept or reject. For this reason they should, in V. Pareto's terms, be regarded as "derivations" and not "residues." Indeed, in constructing normative patterns it is very desirable to draw a distinction between beliefs or residues as deeply unconsciously held sentiments of a people which powerfully motivate their behavior and derivations—the overt statements which these same people use to justify their behavior. This distinction is either explicit or implicit in the writings of many sociologists. By the terms *mores*, *beliefs*, or *internalized values*, subsequently used, will be meant Pareto's residues. My major concern is not with these, but with explicit and overt normative statements which have a bearing on educational research that may or may not be closely correlated with beliefs or mores.

A second distinction stems from theories of human behavior which suggest that residues motivate behavior more powerfully than normative statements. It was made clear in Gunnar Myrdal's *An American Dilemma* in which he contrasts the "higher" and "lower" valuations held simultaneously by most individuals. The higher valuations express their most general aspirations and indicate how they think they "ought to behave." According to Myrdal, however, the lower valuations which are based upon economic, social-class, and political considerations motivate individuals more strongly.

For these reasons the behavior exhibited by individuals is frequently at variance with their higher valuations, i.e., their higher and lower valuations are incompatible. Furthermore, if their internalized value system differs from their higher valuations, their behavior will seem to bear little relation to their own explanations of it. These distinctions are important because the principles which inform research are frequently "higher valuations" and because the measurement of "derivations" is much less difficult than the measurement of "residues."

These distinctions also bear on working definitions of culture and society, both of which are manifest in the kinds of artifacts already mentioned. Here, cultural elements are taken to constitute the approved artifacts of a total society. In particular, the higher valuations—statements of what is or ought to be the best in society—fall into the category of culture. These may be of more or less remote or of recent origin. High fashion reflects currently approved culture. "Pop" culture as a present-day phenomenon may have to be compared with classical culture. Consequently, an unambiguous definition of culture as compared to society is not easy, but since it is not central to the argument the distinction will be left rather vague. Some definition of what is meant by "approved statements" will emerge during the course of the discussion.

In summary, then, our concern is with systems of statements which for various reasons are approved (not necessarily by the majority) as representing higher valuations used by individuals to indicate how they and others ought to behave and how institutions ought to operate. Such restrictions help to simplify the task of establishing theoretical constructs which draw systems of statements together into a coherent and consistent pattern. The arbitrary features of such constructs must be admitted. For the inductivist this would be a criticism fatal to the whole enterprise, but for the hypothetico-deductivist with leanings towards instrumentalism and pragmatism it is not the correspondence with truth or reality which justifies theoretical constructs of this kind, but the uses to which they can be put. These include identifying debates in society and detecting inconsistencies and incoherence in culture. Their roles in analyzing, manipulating, and manufacturing facts in educational research and their contribution to the construction of tests, the interpretation of results, and the identification of possible outcomes are particularly relevant. Nevertheless, the tests of a good construct are not simply instrumental. Internal criteria should also be applied in choosing the source of and establishing a theoretical construct.

Consequently, the criterion of usefulness has internal logical and practical dimensions. Similarly, the selection of statements for inclusion in a theoretical construct depends upon philosophical and observational considerations. They should be comprehensive in the sense that they include the major educational concepts and provide a rationale for the major educational institutions.

For these reasons three criteria of selection are proposed. The first set of statements refers to the grounds for and methods of acquiring knowledge. That is to say, they are epistemological theories. These are no more arbitrary than any other theory. Like any other, they should be capable of explaining data. It is not necessary that there should be universal agreement on any epistemological issue or that there be only one possible stand irrespective of problem and purpose. Alternatives exist and comparative studies reveal national styles of thinking which reflect different epistemological assumptions. Who can doubt that representative Frenchmen put their case in a different way from the spokesmen of other nations? Can it be seriously questioned that Americans habitually argue from pragmatic premises? Or that the English have somewhat unique modes of discourse? In other words, today what constitutes knowledge, how it can be acquired, and the forms of discourse which derive from it show national differences. Comparative studies highlight these differences and make it possible to more successfully compare aspects of educational systems.

Important among these are the curriculum and methods of teaching. Theories of what constitutes knowledge provide a rationale for the selection of content. How it is acquired justifies the way in which selected content is organized in the curriculum. It also implies how teaching and learning best take place. These and other educational institutions can be analyzed by looking at relations between epistemological theories and the institutions through which education takes place.

The second relevant cluster of theories relates to the characteristics of individuals—what abilities they possess, how these are distributed, what motivates individuals to behave as they do, what attitudes they ought to possess, and what constitutes deviation from accepted attitudes and behavior. In short, this is a cluster of psychological theories. Some theories assume that individuals develop in accordance with their individual abilities in a similar way and at roughly the same rate whatever the cultural and societal milieus in which they are brought up. An alternative system of theories maintains that the environment plays a highly significant role in the development of individual children. Within these two broad categories there are many different positions. Again, the introspective theories of Freud, Jung, and Adler may be contrasted with those of the behaviorist physiologist-psychologists such as Pavlov, Watson, and Skinner. There are also psychometricists. Each group takes its particular stand on any issue. A theoretical construct is no less useful because it does not incorporate all the divergent points of view—indeed, because it represents "ideal types" of argument, it cannot.

Nor are such constructs without value in cross-national research simply because they do not themselves take into account every national style in psychology. As in the case of epistemologies, claims are made for some psychological theories that they apply universally. That such theories are not acceptable to all psychologists is evident, and neither are the ways in which data are manipulated and measurements made on the basis of a particular theory. Whether any one of them, as a set of hypotheses about human nature, is subject to decisive refutation is doubtful. Their value lies, as in the case of competing theories or models in the natural sciences, in their role in explaining phenomena and predicting personal behavior. Psychological theories also relate very directly to examination and testing techniques, to policies of grouping for the purposes of instruction, to appropriate teaching aids, and to other aspects of teaching method.

The cultural bias of the third cluster of statements is more readily and widely accepted since theories about the good and just society are much more likely—than is the case with epistemologies and psychological theories—to be regarded as statements of what ought to be the case rather than as referring to what is given. Many answers have been given to questions such as: How should society be governed? What hierarchical relationships should there be between the rulers and the governed? Who should constitute the elite or elites? How should these change? What possibilities should there be for members of the masses to enter the elite? How should they be selected? And on what basis? Terms such as *democracy*, *authority*, *liberty*, *equality*, *stability*, *change*, and many others are debated in numerous political science treatises and dealt with by philosophers and sociologists. Who can doubt that these terms frequently have uniquely national meaning?

It is evident that such theories are used to justify selection or nonselection in a school system. Systems of authority, discipline, ritual, and so on, should be viewed against political science and sociological theories. Methods of controlling and financing schools are justified by similar theories. That is to say, conceptions of society relate both to the internal and external organization of schools in that society.

Evidently, a rationale for any particular aspect of an educational system can be found in one, two, or all three clusters of theory. The practical advantages of finding justification in epistemology, psychology, and sociology or political science should be obvious. Indeed, when such comprehensive support is not available, a school system may well be faced with more or less serious problems. In practical terms, too, when there is a lack of coherence or consistency between the theories accepted by educationists, problems are bound to arise. To identify these problems and those associated with the operation of institutions, it is necessary that the theoretical construct should meet standards of coherence and consistency. For this reason, the source of such constructs should be philosophical. Where can such sources be found and how can the choice of them be justified against other possibilities (an indication that it is not entirely arbitrary)?

The choice of sources for theoretical constructs on criteria of coherence and comprehensiveness implies judgment. The research worker should be at liberty to select the source which seems most appropriate to his purpose. But if the choice is not to be entirely arbitrary, he would do well to consider the way in which philosophers assess coherence. Furthermore, it would be well to select a philosopher whose views are widely debated or inform present climates of opinion. The first test of coherence is therefore logical; the second is instrumental. The application of one type of test may lead to the choice of a different philosopher from the one who seems best able to meet the conditions of the second type of test. Selection is inevitable. Thus for Europe as a whole should it be Plato or Aristotle? In terms of the extent to which their views are still debated, there is perhaps little to choose between them. On the grounds of coherence and perhaps simplicity, I would choose Plato rather than Aristotle. Would philosophers? Certainly for the purpose of cross-cultural research (between Europe and other cultural areas) a useful construct could be based on the writings of either of these Greek philosophers. But Plato's *Republic* probably most fully meets all the criteria. Against such a construct national differences within Europe can be compared. Outside Europe pragmatism has links with the Sophists and consequently beliefs in the United States can be contrasted with those derived from a Platonic construct. Should an American construct be drawn from C. S. Peirce, William James, or John Dewey? Which author measures up most satisfactorily to the logical and instrumental criteria? On the basis of such tests decisions may be made.

Similar considerations hold when constructing models for other main cultural regions of the world. Choosing from among several Hindu, Buddhist, or Moslem traditions presents problems. This is particularly so where no one major text can be pointed to as the most comprehensive and coherent of the literature from a school of philosophy. Nevertheless, from such religious sources as the Bible and the Koran, patterns of theories about knowledge, individuals, and society can be, and have been, derived. Certainly, they offer a basis of understanding the valuations of large groups of people. On both tests it seems legitimate to draw theoretical or logical constructs out of religious literature.

Differences of outlook and belief within a cultural region, e.g., Europe, Latin America, Southeast Asia, imply that national constructs can and should be devised for some types of comparative research. Indeed, in all these areas the differences between nations are considerable, though not perhaps as great as between cultural areas. From among the representatives of a national culture, which philosopher meets the criteria previously mentioned? Unless it is to be completely arbitrary, for England Locke would be a better choice than Shakespeare. For, apart from the latter's lack of commitment to a coherent set of beliefs, which one piece of his writing is the most useful? Of course, J. S. Mill might be preferred to Locke, and his writings would serve to provide a sound model against which present debates could be understood, and on the basis of theories within it research techniques in the social sciences are still devised.

The same might be said about Descartes and Condorcet for France. Both wrote comprehensive treatises which possess an internal logic and coherence. Perhaps Descartes was the more original and certainly many nineteenth-century commentators claimed that France was Cartesian. Both were undoubtedly influential and helped to form French thinking and her system of education. Could a contemporary, let us say, Jean Paul Sartre, be used? Certainly his writings may meet the criteria of coherence and comprehensiveness. His views are very much debated. Some of them are novel; others less so. If against them it is possible to penetrate to the residues which strongly motivate French behavior or to see clearly the higher valuations to which many Frenchmen overtly ascribe, then there are strong reasons for accepting Sartre. On balance a historical figure may meet more of the requirements than a contemporary one.

As for the United States, the acknowledged author of the Declaration of Independence and the Bill of Rights in the Constitution would possess strong claims. Thomas Jefferson cannot be ignored as a possible source for a model of the "American Dream." He was not a pragmatist. Some people would prefer to regard Dewey as the philosopher who most obviously provides a rationale for American education and way of life. It is worth noting that Philip Wiener's opinion is that the early pragmatists were in their own spheres of endeavor, e.g., law, education, politics, psychology, and logic, trying to find philosophical justification for the new practices they saw all around them. In other words, they were seeking to provide coherence in the face of rapid technological change.

The example of the USSR is instructive. Until recently few official announcements of the Soviet Union failed to claim to be based on the authority of Marx or Lenin. Most scholars turn to Marx and Lenin as the source of their most general theories. These determine the kind of research, including educational research on individuals. Regulating-theories drawn from Pavlov, Makarenko, and Krupskaya, which are logically consistent with the more general system of statements, direct much of the research. In countries where cultural pluralism is not officially encouraged, the choice of author and writings presents fewer alternatives than where it is encouraged. Moreover, in these countries, the way in which conceptions of society and culture do influence research on individuals can readily be seen.

One assumption is that most educational research is based upon European traditions. Theories of intelligence which were developed in Europe inform a great deal of the work. Consequently, among the most needed theoretical constructs are those applicable to European contexts. In summary, clusters of statements or theories about knowledge, individuals, and society, and the logical relationships which should exist between them representative of a common European heritage can be derived from Plato. Doubtless, many philosophers would wish to add elements from the Judaeo-Christian tradition. In some spheres of life Plato's views have been decisively rejected, at least overtly, in Europe by all but a small minority of people. Few supporters of de jure political inequality would be found. Residues promoting de facto inequality may well persist. Much less firmly or universally rejected, however, are his concepts of psychological and educational inequality. The result is a lack of logical consistency or coherence in the normative patterns of European models. Such asynchronous changes measured against an ideal-type normative model are the source of identifiable social and educational problems.

Differential changes within and between nations can be measured against a common Platonic ideal construct. The attacks on Plato or his followers were often overt. From these, national theoretical constructs may be derived which will show varying degrees of deviation from the original. The writings of Descartes (Condorcet), Locke (Mill), Hegel (Fichte), Marx (Lenin), and Jefferson (Dewey) are possible sources of constructs for France, England, Germany, USSR, and the United States. On the basis of conceptual analysis, comparisons can be made of key terms (in translation) such as (a) *democracy*, *authority*, *freedom*, *liberty*, *equality*, (b) *personality*, *intelligence*, *ability*, *needs*, *interests*, and (c) *truth*, *opinion*, *rationality*, *objectivity*, *experimentation*, and so on. Within the context of each national system of theories these terms take on a somewhat different meaning. Operational definitions in the form of descriptions of how institutions work would help to sustain and possibly improve comparisons of terms (theories), institutions, and relationships. Simple comparisons by juxtaposition of meaning have themselves little meaning or value unless one national system is judged to be better than the others or worth copying.

Theoretical constructs of this kind are valuable not only as a classificatory system, but also in each of the stages of hypothetico-deductive methods of inquiry. The methods facilitate the identification and intellectualization of certain kinds of problems on which research may then be conducted. Possible solutions to any problem can usually be found among the beliefs and values either of a cultural region or a nation. These may then be formulated as hypotheses which may be tested by comparing predicted events with those which actually occur. A general statement (hypothesis, law, solution) is needed and stated initial conditions known in order to predict outcomes. An important part of these initial conditions are the values (derivations and residues) of individuals in the society. The theoretical construct enables some of the important values, norms, etc., of a nation to be taken into account when nonexperimental (or, when possible, experimental) tests based on predictions are performed.

To be sure, advocates of the "problem approach" would argue that particular norms become important in the light of problems which initially are vaguely perceived. So rising national demands for equal provisions in education may not be matched by changes in conceptions of psychological equality or inequality. Some theories as contextual factors should be given more weight than others. Conceptual analysis of key terms is required. In a cross-national study it would (if the notion of constructs is accepted) be necessary to analyze the meaning given to "equality" in various systems of theories by, for example, Descartes, Locke, and Marx. These would enable so-called empiricists to establish appropriate operational definitions or measuring instruments.

At the moment little conceptual analysis takes cognizance of the cross-national and comparative dimensions of research. Model-making, of course, demands conceptual analysis. Perhaps in empirical research on individuals many of the models are taken for granted or are simply taken from those used in previous research in only one country. On the other hand, perhaps not enough philosophical analysis in education is undertaken with the object of refining operational definitions and measuring instruments.

Naturally, it is for the philosopher and particularly the philosopher of education to undertake such analysis, using logical and linguistic techniques available. Such an analysis of theoretically important and complex issues is designed to reduce or eliminate ambiguity on questions arising from common experience. There is considerable consensus among philosophers (in the Anglo-Saxon world, at least) that this should be their most important, if not their only, role. Clearly they have an important task to perform in the design of experimental and nonexperimental testing procedures.

Doubts have been expressed about the ability of philosophers per se to give a compendious overall view of society and man's place in it. They cannot be the creators of a new order. Creative philosophers of this kind are indeed rare. Nevertheless, some critics of "pure analysis," such as L. A. Reid, have put in a plea that philosophers should perform the complementary work of synthesis in order to get some sense of the whole. A simple theoretical construct provides such a view of the whole. Against such a construct the philosopher may well be able to point to inconsistencies and a lack of coherence in a nation's normative pattern (and these inconsistencies may be tested empirically). He may well be able to indicate what changes would give new consistency and coherence to a confused normative pattern. Such proposals are a necessary part of social reform and, insofar as norms give rise to policies, are extremely important in the planned reform of society and education.

Major Cross-National Studies

Two major cross-national (and cross-cultural) researches on individuals exemplify how conceptions of society and culture are involved in psychometric techniques and suggest how more extensive prior conceptual analysis would have improved the quality and added to the value of the research.

The first of these projects was on moral education and was conducted by the Research Institute of Comparative Education and Culture at Kyushu University in Japan. The second was on achievement in mathematics organized by the Council of the International Project for the Evaluation of Educational Achievement (IEA). The former arose from the desire of some Japanese educationists to solve problems created by the rejection of their traditional system of moral education (Shushin) and its replacement, under American influence, by social studies. The Kyushu group compared moral education in England, France, and Germany with that in Japan in order to formulate new policies for their own country. The first IEA project (others were undertaken subsequently) had as one objective the gathering of information regarded as prerequisite to sound educational planning. Indices of productivity in education were looked for. They were found in measurements of achievement in mathematics.

In the Kyushu research, teams of philosophers, historians, experts in methods of teaching, sociologists, and psychologists carried out a number of research studies. One of these was in comparative education. A multidisciplinary group came to Europe, and subsequently its findings were subjected to the most rigorous scrutiny during a long seminar in Japan. Several detailed follow-up interdisciplinary studies were then undertaken in various parts of Japan.

While brilliantly conceived and most competently executed, this research used tests in Europe which had not been subjected to the usual techniques to ensure validity and reliability. Furthermore, the problems of translating Japanese terms into three European languages were considerable—but instructive. Again projective tests were employed which should, of course, have been interpreted against the normative backgrounds of four different societies. However, test information was supplemented by useful data from other sources.

The main aim of the IEA project was to measure and compare educational outcomes in twelve countries, using psychometric techniques. The search for measures of mathematical achievement and attitudes towards mathematics as dependent variables preceded attempts to analyze the influence of hypothetical independent variables. In this cross-national study, measuring instruments were constructed with the overall intention of making them internationally valid. By conducting a pilot study, by developing a machinery for constructing tests and questionnaires, by pretesting the instruments, by trying to ensure that representative samples were tested, and by sending manuals on test administration to national centers, great pains were taken to ensure reliability and content validity.

The battery of tests in the Kyushu research was much more obviously based upon cultural concepts. Items clearly reflected higher valuations and were, in the postwar climate of Japanese opinion, intended to be representative of Christianity, Western democracy, and more generally, Western society. There were in consequence some surprising results, partly because the tests did not reveal, except negatively, hidden residues or mores. Yet the unexpected results evoked a further critical analysis of the concepts on which the test items were based and enabled Japanese educationists to refine their theories of democracy and morality and from these to deduce logically appropriate forms of behavior.

Nevertheless, both pieces of research raise the question whether it is possible to devise internationally valid achievement and attitude tests. If the answer is no, then what value has such cross-national research? One consideration in looking for an answer lies in the practical value of the research. In the case of the Kyushu study, the aim was clear—it was to illuminate a problem of educational reform in Japan, and the data helped Japanese educationists to formulate a more effective new scheme for moral education than would otherwise have been the case. The theoretical value of the study was that it showed the relationships between clusters of theories, operational definitions, measuring instruments, and test results. Similar theoretical interests are evident in the IEA project. In his summary of major findings, Husén's own admission is revealing. He points out that fewer difficulties than anticipated were experienced in constructing measures of the dependent variables (achievement in and attitudes to mathematics). The main obstacle turned out to be the measurement of independent variables. These included aspects of the educational system and socioeconomic factors, such as family background, for which specially collected descriptive data were obtained; but they were "reported" and not the result of direct observations. National reporters were asked to describe characteristics of the environment by responding to leading questions about it. Student, teacher, and school questionnaires secured reported information about students; about the experience, training, and views of teachers; and about school types, practices, and personnel.

In the study, conflicting interpretations of key terms such as *differentiation*, *academic*, *vocational*, and *comprehensive* were given by respondents, showing the need for basing operationally feasible indices relating to the school system on thorough conceptual analysis. Since most of the hypotheses relating to school systems come into the general category of "selective versus comprehensive," the importance of analyzing these particular terms should be apparent.

Equally open to misinterpretation were key terms describing teaching and learning methods. These were placed on a continuum from "discovery" to "drill and rote" methods. The school climate was measured by reference to "teacher-directed," "authoritarian-based," through "inquiry-centered" references to education. None of these terms can be given unambiguous meaning without reference to the clusters of national theories outlined in the preceding section of this chapter. Each term in relationship to other statements of belief will have a somewhat different meaning in each country—and in some cases a very different meaning.

For example, is the same meaning ascribed everywhere to "discovery" methods? Would it be acceptable in all countries to clarify this meaning by reference to the procedures described by Dewey as reflective thinking in *How We Think*? In short, what epistemological assumptions are made when measuring instruments of this kind are drawn up? Or again—into which context of national norms should "authoritarian-based" education be placed to give the term unambiguous meaning? Is its meaning derived from political theory or from a particular epistemology? These examples simply serve to demonstrate how difficult it is to ensure that such tests are internationally valid.

Fewer difficulties were experienced in constructing mathematics achievement tests because, with insignificant exceptions, the symbols of mathematics are international, and semantic problems are thereby reduced. But it is doubtful whether the aims of mathematics teaching for the majority of pupils are the same or even similar everywhere. A useful distinction was, in fact, drawn by the IEA investigators between lower and higher mental processes in mathematics. It is evident that a measure of international agreement was possible as far as the lower processes of skills were concerned (content, skills, techniques). What is much less certain is that international agreement is possible on what constitutes a capacity to analyze problems or to reason creatively in mathematics. I think immediately of the differences between French and English mathematics for the masses.

Great care was indeed taken in the IEA project to ensure the content validity of the international mathematics achievement tests. It is not apparent from Husén's account that the same pains were taken to make quite sure that the tests had construct validity. The organizers were aware of the problem. Mention is made of the different national aims and objectives in mathematics, but the survey of stated aims and objectives and of the discussions between mathematics teachers from different countries hardly satisfies acceptable criteria of what constitutes rigor in conceptual analysis.

These cross-national empirical studies indeed highlight the need for careful prior analysis in educational research on individuals and show what problems arise in its absence. They are less serious in national studies in countries where higher and lower valuations are virtually universally accepted, where the gap between derivations and residues is small, and, finally, where there is consensus on theories of knowledge, society, and individuality—in short, where there are few normative problems. Today, however, cultural pluralism within nations is increasingly common. Over a wide range of issues there are differences in outlook between political parties, between the young generation and the old, between teachers and the taught, between traditionalists and modernists, and between the privileged and the disadvantaged. It is unwise to assume that unambiguous meaning can be given to controversial terms. Ideal or theoretical constructs of the kind described earlier should not be designed to deny pluralism but to make sense of it. They should help (a) to make explicit concepts of society and culture which are being debated, (b) to assess empirically degrees of consensus and controversy, (c) to indicate which theories lie behind educational measurement, and (d) to throw up suggestions as to the kinds of relationships which ought to be investigated. In both the Kyushu and IEA studies more use could have been made of theoretical constructs; as it was, they were implicit rather than explicit.

Finally, a comparison of the methodological assumptions on which these major cross-national studies were based is instructive. In neither case does the literature available in English make clear what they were. On balance the Kyushu team tended towards hypothetico-deductive procedures. They started with a practical problem, looked for realistic alternative solutions by making comparative studies, and then applied one of several alternatives in the specific context of Japan. They were seeking policies which when applied in practice would bring about desirable changes in society and thus achieve stated aims.

On the other hand, the assumptions behind the methodology adopted in the IEA project stemmed from theories of induction. No urgent practical problem led to the study of mathematical achievement or of attitudes to mathematics. The theoretical problem was to find indices on the basis of which educational productivity could be measured. The hope—a highly laudable one—was that the exercise would lead to an improvement in comparative education techniques. There was, in fact, no real attempt to intellectualize the problems of mathematics teaching or the complex relationships between mathematics achievement and technological and scientific progress. This failure constitutes, in my opinion, one of the most serious weaknesses of the IEA project. Since it was not designed to find and test solutions to practical problems in education, it had less impact on harassed ministry of education officials in the countries involved than the council had hoped.

What assumptions were made is not clear from Husén's account. There is reference to the amount of information which should be collected about school systems and their infrastructures if comparisons are to be applied to "inductive logic." At the same time, much information was collected in the light of clusters of hypotheses connecting mathematical achievement with features of school systems (some of them have been mentioned) and socioeconomic factors. Statistical procedures enabled relationships to be established. The investigators did not deduce a variety of outcomes from particular methods of teaching mathematics or from alternative syllabi and then compare the predicted outcomes with measured achievements. The forms of testing hypotheses in this study are more reminiscent of induction than of the hypothetico-deductive method.

And yet it is evident that technological changes and advances in science have created new manpower demands. Among these one is for trained mathematicians. But is the mathematics needed by research mathematicians and by high-level technologists and scientists the same as that which will contribute to the general education of all young people? If manpower needs are emphasized, some lower and some higher mathematical processes may be required; if mathematics as part of the general education of the masses receives emphasis in educational planning, other processes may be more appropriate. The IEA findings threw useful light on this dilemma. Rightly no ideal mathematics syllabus was offered as a panacea, but it is a pity that more features of the problems associated with mathematics teaching were not made explicit as a preamble to the testing and gathering of descriptive data.

Had such proposals been made it would have been necessary—if the research was to be scientific—to test out ideal syllabi and methods of teaching in selected national contexts. Would the adoption in the United States of an English GCE (General Comprehensive Examination) mathematics syllabus raise levels of achievement in that country? Would the adoption by English schools of Japanese syllabi and methods have the same influence on achievement levels? Immediately these questions are asked, the relevance of the national aims of mathematics teaching, the time devoted to the study by all children, and up to what age become relevant. Successful testing in national contexts would demand longitudinal research. Without such testing, cultural borrowing remains a matter of guesswork, and this is doubtless why practical educationists are suspicious of cross-national studies. The value of such studies, at the moment, is that they illuminate the problems of research on individuals, and they will undoubtedly help workers in the field to improve comparative techniques of investigation and to develop more satisfactory models for cross-national research.

In this article I have stressed that if empirical studies are to be scientific and of practical value, empirical measurement presupposes theories. The latter help us to analyze, manipulate, and even manufacture "facts." The important theories in educational research are normative in that they imply conceptions of culture and society. An integral part of empirical research is, in my view, prior conceptual analysis of these conceptions and theories. Without this we are charlatans—not scientists.

Section III

DEVELOPING NEW MODELS

CHAPTER IX

The Life Uses of Schooling as a Field for Research

HARRY S. BROUDY

The theme of this essay has three aspects which it will be difficult to keep separate:

1. Educational research is in the firm grip of its own orthodoxy, viz., only processes and phenomena that can be translated into publicly observable terms are legitimate targets for research. I shall refer to this orthodoxy as the "dogma of behavioral objectives" and will try to show that it is in large part responsible for the embarrassing discrepancy between the mountainous labor of educational research and its mouselike results.

2. The dogma of behavioral objectives restricts educational research to what I shall refer to as the replicative use of school learnings, whereas the most important life-uses of such learnings for the citizen, or man in his nonspecialist role, are the associative and interpretive ones.

3. Can research be undertaken on the associative and interpretive life-uses of schooling? Can the nonexplicit, i.e., tacit, element in these uses of schooling be studied without self-contradiction?

Research and the Life Uses of School Learnings

Current demands that schools be made accountable for the moneys they spend raise anew the questions: What outcomes of schooling are measurable? For which outcomes is the school responsible? and, Which outcomes are really worth measuring?

Thorndike's dictum to the effect that whatever exists, exists in a measurable amount assured researchers that the search for quantifiable units in education was theoretically sound, but it also relieved them indirectly of the burden of inquiring into what could not be measured. It was difficult to refrain from inferring that what could not be measured did not exist and therefore was not worth teaching or researching.

Up to a point the citizen agrees with Thorndike that the results of instruction should be measurable, and he would agree that certain measurable school outcomes are important. Reading, computation, handwriting skills, and knowledge in the basic fields of study, as measured by standardized achievement tests, would be included in these measurable and desirable outcomes of instruction. The objections to proposed national assessment programs come from the schoolmen, who are aware of the hazards of inferring the responsibility of the school from the performance of even a fair sample of its "products."

These objections do not come from the citizen; his objections are likely to be of another sort. As he surveys the endless variety of social evils, he wonders whether success in school really contributes to success in life. After all, he has been led to believe that school learnings would make their recipients more apt to maximize the social, moral, intellectual, and aesthetic values. And he would like some "objective" evidence that general education is bringing this about—that longer stays in school are justified. Yet on this really important question educational research is faced with an anomaly. It finds itself unable to match life success with school inputs. Some of the school inputs do function in life as they were learned, especially at the skill level, but in the realm of interpretation, understanding, the weighing of values and alternatives, the operations of mind cannot be matched with the school learnings the individual can replicate. This anomaly will be discussed in more detail later. That school inputs function in life not as learned but rather as cues that are used tacitly by the person in facing a life task is the possibility that I believe ought to be explored, if these anomalies are to be avoided. But to advance this hypothesis is to say that the life-uses of schooling are not all replicative but rather of another sort or sorts, and these other uses must now be characterized.

The Uses of Schooling

The pervasive assumption underlying curriculum construction, testing, and learning research is that only school inputs that can be replicated can properly be said to have been learned, and that inputs which cannot be replicated need not and perhaps should not be taught.

The dogma of behavioral objectives is cherished in large part because it makes this assumption operative. The question: What do you want the child to learn? is translated into some observable behavior. Thus, if the goal is good citizenship, this is broken down into such behaviors as reading about elections, voting, thinking so many times a day about the welfare of the community, and manifesting this thought in some sort of talking or doing, and so forth. Now the most obvious next step is to incorporate these behaviors into the curriculum. The closer the match between the life act and the school task, the easier it is (a) to prove that the school succeeded —if the expected life act occurs, and (b) to exculpate the school— if the expected life act does not occur. This, of course, is the doctrine of transfer by identical elements. I wish to deny not only the Thorndike doctrine as an adequate account of what we mean by "transfer," but also a number of its implications and many of its consequences.

There are several ways that a person (P) uses learnings (L) to respond to situations (S). The *replicative* is one in which P instates L, as learned in S, for example, in response to $7 + 6 = ?$ or, Columbus discovered America when?

For P to understand S is to order it to a larger or more general context, for it is not sufficient to understand S merely to utter L items. *Interpretively*, P uses L to classify S as a member of some species, or to hypothesize that S is the cause or result of certain events, or to recognize S as falling under some law or principle. In short, P, upon encountering S, interprets it by constructing some context into which it fits or seems to belong. This context can be a logical or causal or axiological or aesthetic network. We then say that S has become more intelligible or apprehensible to P. For example, if P encounters the statement: "In times of depression spending in the public sector has to be accelerated," he might interpret it by providing a context drawn from Keynesian economics. How much of the Keynesian theory would one have to be able to recall to do this?

Suppose P is asked to use L to *appreciate* S, when S is a poem or a picture or the sufferings of his fellowman in ghettos and concentration camps. To be sure, some replication of previous experience will occur, but primarily what is called for on the part of P is the instatement of a mode of perception and feeling accompanied by imaginative reconstruction. Because virtually all aesthetic expression is indirect and metaphorical, P has to call on L for images and meanings that can be *associated* relevantly (i.e., with aesthetic plausibility) to S. In aesthetic interpretation one uses school learnings replicatively and associatively as subordinate cues within a larger perceptive-imaginative framework.[1] The replication and association are not random, but neither can one predict from the framework just which items will be associated and reinstated. If this were possible, art would be a form of science, or at least a form of direct communication. The received doctrine—that all school learnings to count as learnings must be used replicatively—understandably makes a shambles of all attempts to justify aesthetic education.

When P is called upon to *apply* L to S, then S is an existential situation that is to be changed, and L is the store of facts, laws, theories, and skills presumably learned in school. I think this is an appropriate although not the usual construal of "apply," because we need to differentiate changing a situation by means of knowledge from making a situation more intelligible by means of knowledge. And this is so because to change a situation entails a high degree of familiarity with particulars that understanding does not. I understand the carburetor of my automobile but I lack the tools and dexterity to change it; my mechanic can change it but I am not sure he always understands it.

The familiarity needed for application, I would contend, only the specialist achieves, for only the specialist knows how to perform the operations that have been devised to translate theory into practice. If I am right, application is not the way general education is ordinarily used.[2]

---

[1] For the role of tacit knowing in aesthetic education, see my chapter in *Aesthetic Concepts and Education*, ed. Ralph A. Smith (Urbana, Ill.: University of Illinois Press, 1970).

[2] The description and justification for these uses are set forth in some detail in Broudy, B.O. Smith, and J. R. Burnett, *Democracy and Excellence in American Secondary Education* (Chicago: Rand McNally & Co., 1964).

The differentiations among the uses of schooling do not deny the fact that in virtually every one of the uses the others are involved. I do not need to belabor this relatedness, but perhaps several observations need to be made to anticipate legitimate objections.

1. When more than one type of use is involved, they are not always (perhaps never) coordinate. For example, it may be argued that when I recall the answer to $7 + 6 = ?$ I am also using previous learnings associatively and interpretively. True, but my associations and my selection are guided by the need to recollect a specific answer. Michael Polanyi would call the answer "13" *focal*, for which the associations and interpretations were tacit clues. However, when I estimate that buying both article A, which costs $7.00, and article B, which costs $6.00, is more than my budget can stand, "13" is a clue used *tacitly* in order to comprehend *focally* the import of the budgetary situation. In this circumstance I may not explicitly add $7 + 6$ although I may well be aware of their sum. Should the sums become large, I would have to make their addition an explicit operation, for which the individual number facts learned long ago would be used as tacit clues.

2. The applicative use of knowledge is not always involved in the others but always includes all the others. To apply L means to make focal a very complex and fairly comprehensive context. This, in turn, can break up from time to time into smaller ones. For example, if I were to apply the knowledge that living organisms can be killed by radiation to the task of ridding the lawn of insects, a number of subcomplexes would have to be temporary foci for my attention. Could a gasoline engine such as powers a lawn mower generate such radiation? Could such radiation be so controlled as not to injure the operator or birds?

3. The associative use is always present in the other uses but perhaps only in the arts and in certain types of analytic therapy is it made an end in itself. Aesthetic associations can be controlled by art just as cognitive associations are guided by science, but obviously the control is not of the same sort or degree of explicitness. One might, of course, construe aesthetic judgment as a type of interpretation of associated images.

Accordingly, although the diverse uses do involve each other, the involvement is neither random nor coordinate. The point one needs to stress is that the type of use is determined not so much by the nature of the item as by the role it is asked to play, and one item can play more than one role—but, if Polanyi is correct, not at the same time. Rather, one type dominates the others, using them as cues *with* which it understands the task. The pedagogical significance of the diversity lies in the fact that the interpretive and applicative uses are not learned automatically as a result of mastering the replicative one. And yet the school input, to a large extent, must consist of items that although learned by rote are for interpretive and applicative use. What happens to transform items so learned into interpretive cues?

This, of course, takes us into the domains of the psychology of learning and thought, which I have not the competence to discuss except to venture the guess that the principles of gestalt psychology seem more congenial to phenomena of this sort than do the more conventional associationistic and behavioristic theories. At this level of the problem, it seems as if educational research would have to rely on its parent discipline—psychology—for these more general theories. However, educational research could be interested in those factors that could overcome a reader's failure to understand a passage such as the following:

> During this century, in the United States, about 75 million kilograms of mercury have been consumed; little information is available on its final disposition or on the concentration of the element at specific points in the environment. Recently, it has become clear that compounds of mercury present a substantial hazard. Of particular significance is methyl mercury, a highly toxic substance that causes neurological damage, produces chromosomal aberrations and has teratogenic effects. It is mainly in this form that mercury is found in food fishes. Recent studies have elucidated some of the steps in the accumulation. (*Science*, July 17, 1970)

What concepts and what contexts in the reader are presupposed by this passage? Should the average high school graduate be expected to understand it? How much of the understanding is based on ability of the reader to recall items he learned in school? The same sorts of questions might be raised with the following item:

> WASHINGTON—The House Ways and Means Committee last week approved the most extensive rewriting of the nation's foreign

trade law in at least eight years and possibly in the whole postwar period—a rewriting weighted heavily on the side of protectionism.

Centrality of Interpretative Use

It is difficult to state briefly the argument for the centrality of the interpretive use of schooling, but it may be sufficient to ask: When we read, discuss, choose, how are we using school inputs? And what inputs do we need to have learned in order to cope with these situations? I am suggesting that rarely is the task (S) that of recalling what one learned in school as one learned it. One is rarely asked: How much is 7 + 6? although doing one's income tax may involve reinstating this bit of knowledge. How do you spell "separate"? is probably more typical of the simple replicative use of schooling because in English spelling no other form of learning seems to be feasible.

Nevertheless, the received doctrine is that the replicative use is the typical, almost paradigmatic, use of school learnings, and it is usually averred that what is replicated is a learned connection between S and R; that the more the new situation resembles the learned one, the more likely will R be reinstated. Our achievement tests are constructed on this principle. However subtle the tests become, however varied the test situation, the goal is to check whether the respondent can reinstate an item by recall or recognition.

My argument thus far is that the central life-use of schooling is interpretive, not replicative. It is now necessary to indicate that the interpretive use of schooling involves using school input that once was learned focally and explicitly as clues that operate tacitly in order to further the understanding of a life task. One indirect way of getting at the plausibility of this hypothesis is to note some of the anomalies that result if we insist that all uses of schooling, and especially the life-uses, are essentially replicative, i.e., as the reinstatement of school inputs as put in.

Some Anomalies and Paradoxes

In the early forties a test was designed to measure the public's knowledge of American history; it resembled tests that might reasonably have been given at the end of a secondary school course in that subject, including not a few items involving dates, boundaries, and personages. The American public failed the test so badly that a hue and cry was raised to teach more American history or to teach it better. The American Historical Association (by way of one of its officers) was incensed at the observation of one writer that only idiot savants would remember the sorts of things asked for unless they had overlearned it by use in their nonschool life, and that the results did not prove that the American public would be ahistorical in its interpretation of situations to which history was relevant. Suppose, for example, that the situation S to be understood is the participation of the United States in World War I. It is difficult to imagine that the study of American history would make no difference in this interpretation, even though much of what had been studied could not be explicitly recalled in the sort of detail demanded by an end-of-course examination. Nor would one expect a physician who had never studied biology or chemistry to read articles in medical journals that presupposed some knowledge of these disciplines, yet it is equally hard to believe that a physician who could read such articles would do very well on end-of-course examinations in these areas.

As another odd consequence, consider our schizophrenia about general education. On the one hand we invest much time and money in it, lauding it as a necessary condition for the good life. On the other hand, so much of what is studied as general education cannot be identified in nonschool tasks that it is always on the defensive against encroachments of specialization. Students, whether in law, medicine, or education, are impatient to practice what in life they expect to do and are impatient with having to study anything else. Basic science, as well as the humanities, is under this pressure, and in the high school as well as the college.

Although this pressure is exerted in all professional schools, we are more familiar with it in the training of teachers. One of the largest producers of tests in this country was told at a recent meeting of one of the advisory boards on the construction of tests for teachers that no pencil-and-paper test correlates significantly with success in teaching. "Knowing that" presumably does not correlate highly with "knowing how" and "knowing about knowing how" apparently is not much better in this respect. From this it was concluded that courses in education, especially courses in principles, foundations, and general methods, were useless.

However, general education, which constitutes the bulk of teacher training curricula, was not dismissed as useless, although it did not correlate very well with teaching success either. The reason for this discrepancy in attitude might have been that the educators were taking for granted that the teacher was replicating the subject matter learned in college in the subject matter taught in the schools, and that the lack of correlation might be owing to the fact that all teachers had learned about the same amount of subject matter so that variations would be too small to account for differences in teaching success. I am not enough of a statistician to discuss these conjectures intelligently, but it seems unlikely that all who matriculate into the teacher training program would have a homogeneous grade-point average in their general studies. Furthermore, if there were no doubt about subject matter functioning in the classroom, would one hear so many anecdotes about "my best teacher never finished high school" or that "she knows too much about her subject to be a good teacher"?

After listening to the experts, it seemed fairly clear that on the replicative criterion no tight case could be made for any theoretical input. It seemed a matter of common sense to believe that if one learned algebra in school, one would replicate this in teaching algebra to one's pupils. As for general education, its value was validated not by showing it as an identifiable ingredient in the teaching act, but rather by the faith that somehow this made teachers into better persons. From this blend of beliefs has emerged a conviction that teacher education should consist only of (a) subject matter to be taught, and (b) practice on the job in teaching it. In other words, teacher education would be reduced to some "knowings that" and some "knowings how," but there would be no room for knowledge to teach with, i.e., for knowledge that would not be transmitted to pupils but which would be used to construct contexts for the teaching-learning transaction. Yet, on the possibility of such knowledge for "teaching with," I believe, rests the validity of the claim for a genuinely professional program of teacher education.

The most candid attempt to found a curriculum on the basis of the replicative use of schooling was illustrated by the Committee on the Economy of Time in Education set up by the NEA in 1911, which recommended that nothing be taught in school which could not be identified in the daily tasks of the general population. And we can remember efforts in the twenties to research the words most commonly used in private correspondence, the errors in spelling most frequently committed, and the type of arithmetic problem encountered in this or that occupation or region. All of this made sense on the assumption that the primary use of schooling is replicative. It makes much less sense if the primary use of schooling, especially of general education and foundational studies in the professions, is interpretive. As is the case with many paradoxes, the first step is to examine the received doctrine.

Understanding as Tacit

To begin with, there are two fairly obvious difficulties with the theory that school learnings are used in life by replication either exclusively or even for the most part.

1. The nonschool task (S) is almost never identical with the (L) tasks mastered in school. One has to select from L that which is relevant to S, yet that it is relevant to this particular situation was probably not one of the items learned in school.

2. It is simply the fact that with time—and not too much time at that—items that are not overlearned cannot be recalled. Not only do citizens forget the details of American history, but physicians forget chemical formulas, and engineers forget the specific gravities of the materials they do not use every day. Most college graduates who got A's in mathematics probably could not solve a quadratic equation five years after entering into a mode of life in which they did not practice solving quadratic equations.

We have several possible interpretations of these phenomena. One is that we can use learnings that cannot at the moment of use be made explicit; another is that what we use constitutes a part of what was learned; still another is that, in use, what was learned is on a logical level different from that on which it was learned. Of these three interpretations the last seems to me the most plausible for the following reasons: Although the change in logical level would be consistent with not being aware of the level on which the input had been learned, it does not entail the ineffability assumption in an extreme form. The second alternative merges with the last if what is retained are principles, laws, general rules. However, I find myself often understanding materials without being able to state explicitly the principles presupposed by such understanding.

One of the best known and most thorough attempts to give a theoretical justification for the interpretation I am accepting is Michael Polanyi's notion of tacit knowing, although he was not the first to explicate it.[3] Baldly stated, tacit knowing is distinguished by Polanyi from focal knowing. When, for example, we view two slightly different pictures of the same object in a stereopticon, the final image is known focally, i.e., we have it in the center of the visual field. What about the two separate pictures? According to Polanyi, while seeing the unified picture we have subsidiary or tacit awareness of the two separate images of which the unified image is the resultant. We note several features of this phenomenon.

First, we cannot simultaneously have focal and tacit awareness of the same object. To concentrate our attention on the individual images is to destroy the unified one; to let single images drift out of the center of our attention may allow us to make the individual images focal and thus regain their duality.

---

[3] According to Cyril Burt, Stout in his *Analytic Psychology* (1896, bk. 1, chap. iv) was the first to set forth clearly the account of this type of knowing. It had been discussed even earlier by Berkeley, Lotze, and Burke. "Personal Knowledge, Art, and the Humanities," *Journal of Aesthetic Education* 2 (April 1969): 29-47.

Polanyi's best known work is probably *Personal Knowledge* (Chicago: University of Chicago Press, 1958; London: Routledge, 1958; and New York: Harper Torchbooks, 1964). Among his other books are *Science, Faith and Society* (Chicago: University of Chicago Press, 1946; Oxford: Oxford University Press, 1946; and Chicago: University of Chicago Press, Pheonix Books, 1964); *The Logic of Liberty* (Chicago: University of Chicago Press, 1959; London: Routledge, 1959); and the Terry Lectures of 1962, *The Tacit Dimension* (New York: Doubleday & Co., 1966, Anchor Books, 1967).

Among the papers bearing on tacit knowing used in the preparation of this chapter are: "Tacit Knowing and Its Bearing on Some Problems of Philosophy," *Review of Modern Physics* 34 (1962): 601-16; "The Logic of Tacit Inference," *Philosophy* 40 (1966): 369-86; "Logic and Psychology," *American Psychologist* 23 (1968): 27-43. For a thorough discussion of Polanyi's theory of knowledge see Marjorie Grene, *The Knower and the Known* (New York: Basic Books, 1966); idem, "The Creative Imagination," *Chemical and Engineering News* 44 (1966): 85-93.

Second, in the act of comprehending an entity by means of its parts or components, the parts achieve a vectorial quality; that is, the bearings of the parts *on* the entity come into focus and thus are used as cues to the meaning of the entity.

Says Polanyi:

> There is a fundamental difference between the way we attend to the pointing finger and its object. We attend to the finger by *following its direction* in order to look at the object. The object *is then at the focus of our attention*, whereas the finger *is not seen* focally, but *as a pointer* to the object. This directive, or vectorial way of attending to the pointing finger, I shall call our subsidiary awareness of the finger.
>
> It is our subsidiary awareness of a thing that endows it with meaning: with a meaning that bears on an object of which we are focally aware. A meaningful relation of a subsidiary to a focal is formed by the action of a person who integrates one to the other, and the relation persists by the fact that the person keeps up this integration.[4]

Third, the laws governing the structure of the parts as they are in themselves, i.e., as individual entities on their own account, are not sufficient to describe the principles of their integration. One is reminded here of the theory of emergents that acquire properties not contained in their constituents.

Fourth, the act of knowing in every case involves an effort at integration by the knower.

I think it will take no extended discussion to show why tacit knowing, if there is such a process, gives a theoretical grounding for the interpretive use of schooling and a means of avoiding anomalies and paradoxes that follow from the assumption that the primary use of school learnings is replicative.

The general formula would read as follows: A pupil (P) uses school learnings (L) as subsidiary cues to interpret the more comprehensive situation (S). If Polanyi is right, then during this act of interpretation we cannot recollect and be aware of L focally, for we are using L tacitly or subsidiarily.

---

[4] Says Polanyi further:

"We make sense of experience by relying on clues of which we are often aware only as pointers of their hidden meaning; this meaning is an aspect of reality which as such can yet reveal itself in an indeterminate range of future discoveries. This is, in fact, my definition of external reality: reality is something that attracts our attention by clues which harass and beguile our minds into getting ever closer to it, and which, since it owes this attractive power to its independent existence, can always manifest itself in still unexpected ways." Quoted in Marjorie Grene, ed., *Knowing and Being: Essays by Michael Polanyi* (Chicago: University of Chicago Press, 1969), pp. 180-81.

Next, school learnings (L) may in and of themselves be entities —propositions in a body of knowledge, facts, relations, etc., but in the act of comprehending S their bearing on S alone is what counts, and it is this vectorial quality that renders L into "fingers" pointing at a meaning other than that which they had as individual entities, or even as parts of another system. But nothing in this doctrine holds that one can dispense with learning (L) in its own nature; in other words, it does not follow that in school we need learn only the bearings of L on S. For, that L can have bearing on S is at least in part owing to the properties and relations it has as its own nature, and if we do not learn the infrastructure of L, e.g., as in chemistry, we may not be able to use L in the superstructure of S, when S is the need to do something about atmospheric pollution. Further, nothing in the theory of tacit knowing forces one to say that one can learn the infrastructure of a discipline solely by practicing the use of its components in a superstructural problem. Indeed, the attempt to circumvent the formal study of a discipline by the study of interdisciplinary problems may itself be an instance of the attempt to make interdisciplinary thinking in life situations a replication of interdisciplinary exercises in school.

However, the laws or principles governing L may not be sufficient for explaining the organizing of L into the meaning of S. For example, the subevents in air pollution are explicable by the principles of chemistry and physics, but these principles are not sufficient to structure the problem of air pollution for our social order, just as neither physics nor chemistry can explain the workings of a printing press or an automobile. The purpose of the maker of the machine is needed to organize the parts into a higher integration; the purposes of society are needed to make sense out of the welter of facts relevant to pollution.

Finally, in every act of interpretation or comprehension or even appreciation, the actor has to supply the integrative drive. His purpose, his vision of reality, his value system, his own life commitment organize L into a bearing on S.

If Polanyi is right, then the school inputs need not be found in the outputs in their original form, and not finding them is no proof that they do not function. I am convinced that in the interpretive use of schooling the inputs function tacitly, subsidiarily, even though when they were learned they were focal, but that if they remained focal they could not be used for interpretation later on.[5]

Research and Tacit Knowing

What does the hypothesis of tacit knowing say to the educational researcher? Does it make schooling and the use of it nonresearchable? Does the assumption that it could be researched involve a self-contradiction, namely, that the tacit would thereby become explicit? Let us take up the last point briefly.

Suppose we succeeded in discovering just what input of a biology or chemistry or literature course *does* function later in life; this could then be taught as it is used. Is this not, therefore, just another way of finding an input that can be replicated as an output? The objection would be more cogent if we could assume that in order to get the desired output, we need only make it an input. This is a bit like saying, "If you want the body to put out energy, equip it with a motor, not food." However, if, as one might suspect, the input is transformed in becoming an effective schema for association and interpretation, then the objection loses much of its force. This transformation takes us into learning theory and research, and I can only think that here, too, simple S-R connectionism has not proved apt in accounting for the processes that apparently are instated when knowledge is used interpretively or associatively. This may be a fruitful link between research into school learning and curriculum.

The key, it would seem, is to get some insight into how materials taught in the classroom become transformed into tacit clues for interpretation of various kinds: scientific, historic, aesthetic, sociological. If I am at all correct in my conjecture that interpretation of an educated sort is accomplished by means of a residual schematism left by the study of an intellectual discipline, then failure to accomplish certain interpretive tasks would be evidence of the absence of an adequate interpretive context. Can we devise tests that would reveal such absences?

---

[5] In this connection it is interesting to note the hypothesis about tranfer arrived at by Hendrix: (1) For generalization of transfer power the unverbalized awareness method of learning is better than a method in which authoritative statement of the generalization comes first. (2) Verbalizing a generalization immediately after discovery may actually decrease transfer power. Gertrude Hendrix, " A New Clue to Transfer of Training," *Elementary School Journal* 48 (1947): 198.

I am of the opinion that we might construct such tests by using samples of articles on politics, economics, art, and science which appear in high-grade metropolitan newspapers and in journals designed for an educated readership. Two questions could be asked of the reader: Are you interested in the topic being discussed? and, Can you understand well enough what has been written to answer questions about it? Case studies on responses of readers might reveal those concepts and terms which seem to be necessary for the understanding of the selections, and one might then study whether or not such concepts were the product of formal study.

The following samples of materials represent what I would characterize as an "educated" level of discourse about topics that are presumed to be of interest to and within the comprehension of people who have a good general education by American standards—i.e., what one has received by way of general education through the fourteenth year of formal schooling.

> Berlioz's *libretto*, based on Books I, II and IV of the *Aeneid*, carries the *mythic* tale in a long arch from the *Trojans' discovery of the horse at their gates*, through the *death of the Trojan women*, into the *Dido and Aeneas* episode at Carthage, and even foreshadows the rise of *Hannibal and the Punic Wars.* Along the way, we continually find the classical past refusing to be written off as irrelevant. Race warfare, for instance, dominates the Carthage episodes (Aeneas wins Dido in the first place by rescuing her from the threat of the black king Iarbas, who marches on Carthage with the intent of marrying the white queen). (*New York Times*, July 26, 1970.)

The italicized passages or words require contexts from Greek and Roman mythology, musicology, and Roman history. The whole paragraph would be virtually unintelligible to anyone who had not become familiar with the concepts out of which these contexts could be constructed. Yet, I suggest that the precise details about these mythological and historical events need not be recalled in order for a sufficiently functional context to be formed.

> WASHINGTON—Now that the nine-month slide in the economy has apparently about touched bottom, how vigorous should the Government be in pushing the upturn toward full employment? This wholly new issue emerged dramatically last week as the Congressional Joint Economic Committee held an important set of hearings on the state of the economy at midyear.
>
> Three viewpoints emerged from entirely respectable sources. Ranging from left to right they were:
>
> (1) Pump the economy up vigorously through the use of both fiscal (budget) policy and Federal Reserve monetary policy (money and credit). This viewpoint was pressed last week by two leading Democratic economists . . .
>
> (2) The time for an expansionary policy has come, but through monetary policy, not fiscal policy. Government spending should be held in check and Congress should raise the revenues, such as higher postal rates, asked by the President, but the Federal Reserve can properly expand the money supply and bank credit faster than has been the case so far this year.
>
> (3) Present fiscal and monetary policies are about right—assuming Congress acts roughly in line with Administration wishes on spending and revenues—and there would be a real danger of wrecking the progress made against inflation if a policy of pumping up the economy were adopted. (Excerpted from the *New York Times*, July 26, 1970.)

Or, to take a somewhat different topic in the same field:

> Hope, Francis Bacon once commented, makes a good breakfast, but a lean supper. As Latin America enters the seventies, her governments tremble beneath the bruising tensions that separate hope from fulfillment.
>
> "Here is a subcontinent," historian Arthur Schlesinger, Jr., observes, "where one-eighth more people than the population of the United States subsist on less than one-eighth of our gross national product, where 5 percent of the people receive a third of the income and 70 percent live in abject poverty, and where in country after country the political and social structures are organized to keep things that way. . . ."
>
> Still, economists know what is required within Latin America to move it into an era of adequate, self-sustaining economic growth. There is general consensus on the necessity for far-reaching agrarian and fiscal reform, for increasing internal savings and enlarging internal markets, for regional economic integration and for more favorable trading arrangements with the developed countries. Most of all, there is the need to bring into the national economic life the large numbers of Latin Americans, amounting in some countries to the greater part of the whole population, who are now, for all practical purposes, subsisting

outside a money economy. (Frank Church, "Gunboat Diplomacy and Colonialist Economics," *Transaction*, June 1970.)

I think it would not be too difficult, by consulting competent men in economics, to identify the concepts needed to understand the import of these articles and to determine which of these concepts are most likely to have been learned in formal courses in economics. It should then be possible to test the hypothesis that the inputs and the outputs do or do not match or the degree to which they do. It might also be possible to design test-tasks in which the testee carries on a discussion with a partner who uses concepts and relationships which may or may not have been studied by the testee.

The same sort of conceptual analysis can be done with the following passages and test-tasks devised to appraise the functionality of schooling that no longer can be replicated as learned.

> Radio astronomy, having coined the words quasar and pulsar, has now added linars to the growing list of strange objects inhabiting the visually dark areas of space. Linars—based on the words line and star—are point sources which emit with extraordinary energy, at wavelengths characteristic of the spectral line of particular chemical compounds.
>
> Radio astronomers in Britain and elsewhere have for several years been studying large diffuse areas in space which emit on the comparatively long wavelengths of hydrogen or of the OH hydroxyl radical, but recent work—carried out mainly in the United States—has shown that sometimes embedded within these large areas are very much smaller sources pumping out energy on much shorter wavelengths. They are the brightest radio objects in the sky.
>
> Successive observations revealed that vast regions of space were far from empty, and the discovery of the characteristic spectral fingerprints of carbon monoxide, hydrogen cyanide, and ammonia suggested that radio astronomers were looking at protostars in the early stages of formation. (*Manchester Guardian*, August 5, 1970.)

> ### *Up the Rebels*
>
> Kafka-like, *The Revolutionary* purports to take place in an unnamed city in some unnamed country; unlike Kafka, the events are all too prosaic, the style pedestrian. The film is based on Hans Konigsberger's novel, which appeared some years back, and what was hypothetical then is painful reality now. The film thus has the look of simply reflecting and abstracting some typical events: student dissent, faculty obtuseness, meetings, activism, symbolic violence. Undoubtedly, the attempt was to universalize the conditions and the types, but this becomes rather laugh-

able when we can clearly recognize the English settings, in the midst of which the police wear uniforms rented from an establishment presumably specializing in comic-opera uniforms.

How much more finely wrought—and not only by comparison—is *The Virgin and the Gypsy*, based on the last work of D. H. Lawrence. Here rebelliousness, rather than rebellion, is the theme, and, specifically, the revolt is against convention, hypocrisy, and not so much the social order as the social niceties. The virgin of the title is a North England girl who returns to her rectory home fresh from a French finishing school. The first, faint stirrings of sexual needs are making her uneasy. Upon which enters the gypsy. The darkly brooding, elemental, and handsome fellow knows a girl in heat when he sees one, and he unleashes her sexual force, but only after a dam breaks and the resultant flood chases them to her bedroom. Hollis Alpert, *Saturday Review*, July 25, 1970.

Two criteria would seem to be relevant to gauging the efficacy of general education for these uses. One is the degree of interest in the samples. I have found through unsystematic means that several years of college work in the more or less usual distribution of general studies courses broaden interests. I have measured this roughly by pre- and posttests on the special sections of *The New York Sunday Times* given to entering freshmen and at the end of the first semester of the senior year. Presumably, course work in economics, history, science, arts does increase interest. Systematic study with carefully controlled groups could refine these conjectures. Further, these tests—which also involved a judgment as to the ability of the student to comprehend the articles indicated as "interesting enough to read in detail"—also showed a marked improvement. Again, more careful research might validate the confidence with which one could attribute the improvement to course work.

Another line of inquiry concerns the school's part in the formation of images, and for this purpose such instruments as the Rorschach inkblot tests could be employed as well as materials from the arts in various media. The imagery required for the appropriate aesthetic response to Milton's poetry, for example, depends heavily on Latin and Greek root words that have been incorporated into the English language. For example, "an elephant endorsed with towers" is unintelligible without the imagery connoted by "endorsed", i.e., put on the back of. The elimination of Latin from the curriculum could be ruinous for the study of some English poetry.

But perhaps the most important general indicator of the functionality of schooling in life is the level of language that the person uses in speech and reading. Size of vocabulary, speed, and level of comprehension are familiar criteria of language usage and are used by the common man as well as by test-makers. That the assessors of general education so often forget to credit schooling for language facility cannot, I think, be attributed to their failure to realize that it is largely the result of schooling for most people, but rather to the belief that verbal facility as such is small potatoes in measuring the real worth of a man.

Here, research can help in determining the role language facility plays in the kind of reading and discussing the citizen is called upon to perform. Much of the text in the few samples cited is opaque or at best translucent not only to readers who have not studied science or economics or history or the arts, but also to those whose general language facility is limited. Leaving aside the factor of speed, the level of abstraction of words, phrases, and sentences, one would think, are crucial to understanding. To what extent does the input of high school and college determine this abstraction level? And how does this input change with the passage of time?

Researchers do not lack the ingenuity or methodological dexterity to probe into these questions. But the researcher should not approach the problem with the firm assumption that what has been taught in school either will function in the life test *as it was learned* in school or will not function at all.

It may turn out that the way an item was learned in school does prevent it from functioning in the life test-task or causes it to function in undesirable ways, e.g., as certain ways of learning history or geography do. Such findings would pave the way for experimentation with other methods of teaching these subjects but not necessarily for teaching them *as used in the life test-task*.

To test fairly the functionality of school learning would require using subjects who had been away from formal schooling long enough so that normal forgetting could have occurred. The case for using one's schooling successfully in nonschool life may

CHAPTER X

A Model For Making And Testing Value Judgments

LAWRENCE G. THOMAS

In the past twenty-five years there has been a remarkable amount of communication between philosophers and social scientists concerning the empirical study of values. Some social scientists, of course, were edgy about this development, insisting that a genuine science of human behavior could have nothing to do with normative preferences and value judgments. Many philosophers also deplored this trend, arguing that only the philosophic methods of logical analysis and disciplined intuition were adequate to deal with value questions. But these two extremes have not prevented naturalistic-minded philosophers from cooperating with behavioral-minded scientists in exploring several dimensions of values.

In one sense value judgments have been objects of empirical study for as long as there have been social sciences. If value judgments are interpreted as expressions of preference or declarations of worth, then wedding rituals, laws of inheritance, and the progressive income tax qualify as value judgments. In fact, most of our social arrangements and traditional culture are formalized expressions of value. Sometimes the focus of attention is on the established content, as when an anthropologist is exploring the religious practices of the Kwakiutl or an economist is describing wage structures in the textile industry. In these instances, the established content consists of valued ways by which certain people channel their interests and reach at least a tolerable form of adjustment. At other times the focus is on the process of change, such as the change in economic values represented in the Industrial Revolution or the change in political policies in the United States after World War II. Values of one sort or another constitute the main subject matter of the social sciences.

But in another sense this approach to value judgments is not especially relevant to the concern of the politician, the labor leader, the business executive, the missionary, and the philosopher. The value content of the social sciences does not answer their urgent questions. The content of these disciplines is drawn from established, habitual value judgments, whereas these leaders are seeking new departures. It deals with innovations only in the descriptive sense, not in deliberative judgments of better or worse. It depicts the process of change and categorizes it, but it does not tell how competently the changes were made. In short, the social sciences provide useful background but do not deal directly with the concerns of the decision-makers.

A philosopher, on the other hand, is profoundly concerned with the problems faced by politicians, labor leaders, business executives, school officials, and the like in making their value judgments. He too is interested in how best to make judgments, not merely in what were the results of past decisions of specific groups. And he is interested in the process of making value judgments, not just to find out how it happens to be done by assorted groups, but to study whether it has been done poorly or well. This shared concern may account for the often reported observation that teachers, administrators, and similarly situated officials make more use of a philosophy (or ideology) than of scientific generalizations in reaching their decisions.

My intent, however, is not to oppose the treatment of value judgments in the social sciences to their treatment by philosophers and policy-makers. This would be a false opposition in any case, for a decision-maker who lacked knowledge of the social sciences relevant to his work would be largely limited to idiosyncratic experience and raw empiricism for the factual basis of his judgments. My intent, rather, is to urge the extension of scientific inquiry into these areas of special concern to the makers of value judgments. As matters now stand, scientific investigation falls short of any direct contribution toward answering the crucial questions, the "should" questions.

Consider the inquiries that could be made about the topic of controversial questions in the classroom:

1. Are controversial questions discussed in the classrooms of this school?
2. What are the effects of permitting the discussion of controversial questions in this school—on the pupils, on the parents, or other teachers, on civic clubs, and so forth?
3. Who opposes the practice? Who favors it?
4. Should controversial questions be discussed in the classrooms of this school?

The first three are straight empirical questions and, with proper qualifications and controls, could be investigated right now with the techniques of the social sciences. The fourth is a value question. It is also the most important of the four for school boards, administrators, and citizen groups, as well as for teachers. But, unfortunately, it is commonly considered unresearchable by scientific procedures. As we shall see later, this negative attitude occurs when the value question is treated quite separately from the preceding three questions instead of being analyzed into elements that could organize and interpret the investigations called for in the first three questions.

The same case could be made concerning the question of establishing a dress code for the school. Empirical inquiries can be made on a number of relevant points—who favor and who oppose a dress code, how many other schools have one, what consequences elsewhere have been noted, the legal rights of pupils and parents in this matter, the legal authority of the superintendent in dealing with these and similar problems. After these questions have been answered, however, the crucial question still remains: Should the school have a dress code?

And so it goes with similar questions of educational policy. The controversial question is not how many schools, for example, use ability grouping of their students but should the school use ability grouping. Not so much the present distribution of students by race in these schools as whether there should be proportional representation by race. Not so much a description of the legal powers of college administrators as how they should treat college students who use organized violence. In short, the most critical questions in education policy are characteristically value judgments, which are seldom considered to be researchable. If these value questions are not considered researchable, how are they answered?

The ways in which these major questions are answered at present are often opportunistic, capricious, or intuitive in the most unsystematic sense of that term. Sometimes the answers are sought in the majority vote of the public, as when school board candidates campaign for popular support of their particular preferences. More frequently, special-interest groups advance their simplistic answers, and school officials are expected to yield to the strongest pressures. Or, in direct contrast, the top administrator is regarded as the boss, and whatever he says goes. Even when expert consultants are called in to help, they are often expected to recommend the right answers instead of providing the facts and alternatives that the decision-makers need to deliberate on in reaching their own judgment. All of these are poor ways to attack such important questions of educational policy.

What is needed? Some would say we need stronger appeals to virtuous behavior, to respect for proper authority, to loyalty toward our moral traditions, to greater deference toward the judgments of scholars, to deeper commitments to universal standards. Helpful as these might be in raising the moral tone of policy controversies, they alone are not equal to the task. We do not need to standardize our values, but we do need to standardize our way of making and testing value judgments, especially if we take these questions seriously and want the judgments to be made competently. To get the maximum attention from empirical researchers, the procedure should be as open as possible to the attitudes and techniques characteristic of scientific inquiry. Nothing less than these requirements seems likely to insure the full cooperation of social scientists and educational philosophers and to enhance the quality of major value judgments in education.

The Proposed Model

The process of making and testing value judgments is conceived here as a modified form of problem-solving in the naturalistic, relativistic tradition of John Dewey and the latter-day experimentalists. It consists of six distinguishable stages somewhat artificially abstracted from the flowing process of actual evaluations. They are presented here in the approximate order in which they usually occur in action, but any stage may be returned to or repeated in actual evaluations. Only one stage is ever likely to be omitted in the actual process, and that is the third one—achieving a common frame of reference, a state of affairs which often exists already in communities or groups that have worked together successfully. After seeing first the distinctive character of each stage only briefly (to get the continuity or flow of the total process) we shall then look at each stage in considerable detail, seeking clarifying illustrations and opportunities for empirical research.

1. *Undergoing conflict or lack.* This phase is the stimulus or occasion for making a value judgment. There would be no motivation to make a judgment, let alone test it, if there were no disruption or interruption of our habitual ways of acting, no forced choices, no sense of lack and hunger, no awakening of a desire for something new that cannot be obtained by the automatic operations of past habits.

2. *Diagnosing the situation for clues.* This phase constitutes the search for information, not merely to locate the possible causes of the disruption or the reasons for the appeal of the new object, but also (and more frequently) to acquire means for planning what might be done next. The data needed include not only facts and empirical evidence but also the experiences and interpretations, the opinions and interests of others.

3. *Constructing a common frame of reference.* If two or more persons are involved in making and testing a value judgment (as is usually the case), they are not likely to have fruitful discussions, debates, and appeals to the evidence unless they share a common perspective on the issues. Clear communication and the achievement of common understanding depend upon sharing the same assumptions, the same facts and values being taken for granted. These are the crucial features of a frame of reference, and agreement on these assumptions may have to be sought before the preceding phase—diagnosing the situation for clues—can be pursued very far.

4. *Deliberating on means-end relationships to reach a firm judgment.* A distinctive feature of this stage is the organization of the facts, opinions, interests, and other information collected into tentative judgments or hypotheses. When these are expressed in the form of means-end relationships, where a tentative end in view not

only determines the desirability of possible means but also finds its own desirability heavily influenced by the character of the available means, the amount of deliberation demanded at this stage is often extensive and exacting. The intellectual product at this point is a value judgment based on prior diagnosis of the situation, seen from a specified frame of reference, and now considered worth testing in practice.

5. *Taking action on the judgment.* This testing stage is needed not only to validate the judgment at hand but also to provide evidence to use in making a better judgment next time.

6. *Reviewing the experienced conequences.* This, like the first stage in this process, is primarily an undergoing stage but revealing at this point the worth of the previous investigating, consulting, deliberating, and choosing. As such, it deserves serious attention and reflection. The enjoyments, difficulties, interests, and factual connections which were once anticipated in the matured value judgment are now compared with the autonomous content of the experienced consequences of acting on that judgment. How well do the results fulfill the anticipations? All in all, do these results resolve the conflict or fulfill the lack that grew out of the initial undergoing in this process? What has been learned that could be used to produce a more accurate and more satisfying judgment another time?

This, in brief outline, is the six-step model proposed to describe a competent performance in making and testing any value judgment. Each step needs a more detailed development of its possibilities, clarified by illustrations. The opportunity for systematic research to refine and improve the tests at each step would then be worth noting. To this expanded treatment of each step let us now turn.

UNDERGOING CONFLICT OR LACK

The initial undergoing provides the impetus for reexamining what we have been doing, but when it is lifted from the total process for closer inspection, we find it is suffused with uncertainties, ambiguities, and dilemmas. The undergoing does not, in itself, provide a purpose; rather, it creates the need to define a purpose.

The easiest way to describe this initial undergoing is to focus on the disruptions, interruptions, and conflicts which are immediately felt. For example, a principal of a high school, accustomed to the conventional middle-class range of attire and grooming, has had his expectations radically disrupted in the past few years. He might be able to tolerate the changes in dress, but when the athletic coach will not let long-haired boys play on the team, when teachers complain that their objections to mini-skirts on girls are answered by a shift to Levis, when the student employment agency cannot place some youth solely because of appearance, and when some parents demand that their children have the right of self-expression in dress, his established habits and attitudes are threatened with disruptions, incongruities, and conflicting demands. He is catapulted into the process of making a new value judgment.

Dewey is usually credited with emphasizing the disruption, the conflict, the dissatisfaction that often launches problem-solving, but he uses the term *lack* as often as he does *conflict.* A lack could be considered a product of "disruption of continuity" or of disappointment with the results of established habits, but this conception tends to limit motivation to the restoration of equilibrium. Human beings, when they are healthy, rested, and unthreatened, want more than the serenity, tranquility, or contentment of equilibrium. When most of their basic deficiency needs have been taken care of (as Abraham Maslow, Carl Rogers, William Livingston, Hadley Cantril, and others have pointed out), human beings naturally seek new value experiences, new adventures, new risks they feel they can control, new refinements of old joys, new worlds to conquer. It seems to me that these can properly be called *experienced lacks* without doing too much violence to the conventional meaning of the term. The range of lacks we experience might be described on a continuum. At one end would fall the lacks arising from boredom, ennui, a vague restlessness, an unidentified unsatisfaction with the tedium of overcontrolled experience. These qualities of a situation can be just as upsetting and problematical as an external invasion or an unexpected interruption or an unasked confrontation with an opposition. They constitute dissatisfactions with the present state of affairs. At the other end would be curiosity, the challenge of the unknown, a taste for variety, the thrill of adventure, the excitement of novelty. When a person is psychologically ready for these appeals, they occur as lacks in his present activities. They represent unsatisfaction more than dissatisfaction, but they upset the equilibrium of present activities just as surely. Moreover, they combine urgency and indeterminacy; they call for doing something different without providing adequate clues or pointing a direction.

Suppose we take as our continuing example a proposal to achieve proportional distribution of races in a district's schools by bussing children beyond their neighborhoods. It is easy to see how this proposal can appear as a threat or possible disruption to established ways of doing things. If the proposal remains as only an external threat, however, no new judgment is required on present policy. The proposal is simply bad, evil, the enemy. Our present ways of allocating pupils to schools are good and must be defended. But if the proposal upsets our confidence in the established ways, if it raises a conflict or reveals a lack in what we always thought we wanted to do, then a new judgment must be sought. Our conflict cannot really be identified at this initial stage of undergoing and becoming upset, but it may turn out to be between a strong preference for neighborhood schooling and a profound belief in fair opportunities for all children. Or the felt lack might become identified later as this question: Will racially proportionate schools make education more real and relevant for my children? In any case, the motive or impulse to action is to remove the conflict or do something about the lack as now felt.

Since the problem is not yet objectively defined, the conflict or lack is not yet located and cognized. This outcome is desired but it comes later. Right now the entire situation is indefinite and problematical to us. If we were being studied by scientific observers, they could not see the character of our conflict or the nature of our lack yet any more than we can (although they might have more hypotheses), but they could surely see our perplexity and uncertainty. If this initial stage out of the entire six-step process were to become itself an object of empirical investigation, the following questions would be appropriate:

1. What are possible ways of locating (or creating) situations in which specified kinds of persons feel an important conflict (or lack) in their established ways of handling such situations? Which ways tend to evoke hostile defenses, which ways tend to evoke

aversion and escape, and which ways usually lead into the second stage of making a new value judgment?

2. What signs or indices of undergoing a conflict (or lack) do various types of persons display in various kinds of situations? Can the dependable signs be categorized for economical use? Can they be predicted under specified conditions?

The answers to these questions would not only enhance our understanding of the beginning of competent value judgments but would also make the deliberate inducing of them more feasible and more reliable.

DIAGNOSING THE SITUATION

Having undergone an upset for unidentified reasons, we must now do something. Doing something is seldom, if ever, at random. It will be guided, no matter how tentatively, by some suspicion, some hypothesis, some hunch, some intimation. And the information found under this tentative direction will, in turn, affect the guiding hunch—confirm it, refine it, alter it, refute it. The search is very likely to be more forward-looking (i.e., focused on what needs to be done) than backward-looking (i.e., finding out how this difficulty occurred), although the difference is a matter of emphasis instead of kind. This difference in emphasis has traditionally characterized the distinction between the concerns of evaluators and the concerns of scientists,[1] even though a value judgment may be significantly improved by basing it on scientific knowledge of what causes the initial difficulties.

Since the purpose of this diagnostic stage is to gather information relevant to interests to be incorporated soon in a value judgment, the range of information sought will extend considerably beyond the usual kinds of hard evidence traditionally favored by educational researchers. The range will include not only official reports, vital statistics, test scores, and persistent attitudes, but also beliefs, hopes, current preferences, past fears, and present suspicions and doubts. A wide variety of people may have to be consulted to learn what interests are actually at stake in this unsettled situation.

The main achievements to be sought at this stage include (a) identifying most of the important interests that are thought to be desired (and functioning as goals, outcomes, criteria, or standards); (b) identifying most of their relationships to possible facts, trends, resources, and facilities that are thought relevant to the achievement or preservation of these important interests; (c) finding evidence of differences in judgment and interpretation that are probably worth investigating which arise from possible differences in frames of reference. The valuational question, for which a value judgment will soon be made and tested, can usually be stated loosely, vaguely, and in several versions at this point.

---

[1] See Peter Caws, *Science and the Theory of Value* (New York: Random House, 1967), chap. 4.

Returning to our example, let us say we are initially guided by a sense of conflict regarding school bussing. To locate the conflict, we seek to identify the important interests involved, our own as well as those of others. If we find that some parents oppose bussing because they chose their homesite to be near the highest quality school and others favor bussing in the belief that white parents will pay more attention to the educational opportunities of black pupils if these latter are attending the same school as their own children, we have tentatively located a conflict between groups of parents (even though drawing a genuine issue here must wait until the fourth stage). If we find a strong commitment to maintaining high educational standards and to providing equal educational opportunities for all children but also find profound doubts about the competence of the teaching staff to do well with pupils of heterogeneous backgrounds, the conflict is between professed ideals and presumed facts. If we find ourselves committed both to providing equal access to the best schools in the district and to avoiding any transfer of pupils from their own neighborhood schools, the conflict may be a disjunction between what Charles Morris calls a "conceived" value and an "operational" value[2] (i.e., between what we profess to be generally desirable and what we actually practice). This kind of conflict calls for self-diagnosis more than for environmental diagnosis.

Depending on which conflict of interests now guides us, we shall then look for such relevant facts as the distribution of races in the different residential zones, the comparative school facilities at predominantly white and predominantly black schools, the amount of parental participation in the program planning of each school, the standards of teacher selection maintained by the school board, the content of the in-service training provided for teachers at different schools, the costs and risks of bussing, and so on. We should recognize, however, that at this stage the relationships between a tentatively identified conflict (or lack) and the relevant information sought and acquired is both reciprocal and unstable, each side of the relationship being capable of altering or redefining the other. But as this diagnostic stage approaches fulfillment, certain valuational questions will emerge as tentative definitions of the problem. For example, how can we secure equal educational opportunity without bussing the children? Or how can we liquidate the white middle-class opposition to bussing?

---

[2] Charles Morris, *Varieties of Human Value* (Chicago: University of Chicago Press, 1956), chap. 1.

If we were to become objects of empirical research at this stage of the judgment-making process, significant contributions to our understanding of the process and to our competence in using it could be realized from answers to such questions as these:

1. How can an observer identify the kinds of interests and contingent facts that different groups of participants actually perceive or acknowledge in their diagnosis? Could he use the counselor's technique of having the participants talk about their perceptions and dilemmas? Would a role-playing technique, where one participant is asked to instruct another, be effective? Would asking them to write down the pros and cons of alternative choices be useful? (Presenting a check-off list of possibilities is probably too "instructive," and the excessively intellectual game of asking them to face imaginary choices with preformed sets of consequences should also be avoided).
2. What evidence will discriminate a genuine involvement of the participants in making diagnoses from a mere willingness to role-play or engage in a game with the observer?

CONSTRUCTING A COMMON FRAME OF REFERENCE

Specific attention to the frame of reference is needed when more than one person is involved in reaching a value judgment and when the persons involved have different backgrounds. Similar frames of reference are usually held by those of similar economic status, educational status, and having similar social and vocational opportunities. The similarity will probably be implicit and unnoticed, but as long as it is shared, it need not be made explicit for study. Hence, when the construction of value judgment is of concern to only a single individual or to a group of similar backgrounds, this third stage can probably be skipped. For all others this stage deserves careful exploration, because so much of what a person sees, believes, or ignores is determined, usually unconsciously, by his frame of reference.

Since the most important features of a frame of reference are the facts and values taken for granted, they are the most difficult to identify. The things that we take for granted are the last things we look at or question. Often we have to be irritated, annoyed, or outraged by the suggestions or acts of others to discover the hidden values and beliefs that we have been taking for granted. Instead of indulging our irritation or anger at this stage, we should be wise to take them as clues to recognizing our own frame of reference. Another useful clue is the qualities we praise and the traits we condemn in others. These are often our ideals for ourselves—and sometimes the suppressed tendencies we hate in ourselves.

After some humbling degree of self-awareness has been obtained on these hidden biases, attention can be given to deliberately constructing a common frame of reference. One of the most useful skills in accomplishing this is facility in detecting possible assumptions underlying the diagnoses of, and proposed solutions to, the difficulties under discussion. For example, in the argument over the bussing of pupils away from neighborhood schools, the claim might be made that this additional bussing would increase school costs 10 percent. Regardless of the factual accuracy of the case, the claim could rest on the assumption of no educational benefit coming from the extra cost (and the implied waste of school funds) or on the assumption of proportionate (or even greater) educational value being derived from this expenditure—and it surely makes a difference which assumption is held. Or consider the claim that the mixing of white and black pupils from different backgrounds will lead to more frustration for blacks and lower-quality schooling for whites. Before plunging into debate on the claim, let us look first for its possible assumptions. Does it assume that educating consists of presenting the same academic materials to all children for mastery? Does it assume that the differing backgrounds of pupils are irrelevant to the proper content of the curriculum? Does it assume that motivation for learning is best set by the amount of competition stimulated among pupils? If not all these plausible assumptions are acceptable, what alternatives exist and how would they affect the meaning and validity of the claim?

There is no guarantee, of course, that this process of exposing assumptions will always result in consensus on a frame of reference, but it certainly increases the prospects of achieving agreement on some assumptions and identifying the genuine disagreements. The prospects of consensus can be enhanced even further if common agreement on some principles of method can be obtained fairly early. For instance, disagreements on *facts* assumed could be resolved by empirical investigation if one of the *values* assumed by all involved is the superiority of publicly tested evidence over one's personal beliefs, no matter how cherished. Another facilitating principle is agreeing to treat all value judgments as hypotheses, no matter how confident we may be in their authoritative source. Especially desirable is some skill in trying out different frames of reference, occupying first one and then another, to get their distinctive views of the world. In this connection the group members might adopt the principle that the degree to which one understands a different set of assumptions is measured by how plausibly they can be stated by one who does not agree with them.

These strategies and techniques are not needed for the deep moral commitments that are shared by virtually everyone growing up in our western culture. And, on the other hand, they would not work to resolve a clash of beliefs or attitudes at the visceral level (e.g., favoring and opposing the close association of blacks and whites as equals). The latter conflicts yield only to time and further experience. But they are suitable for the value assumptions of moderate intensity or depth and of easily acknowledged diversity. To illustrate this level and type, the participants in our example might reach agreement on the following value hierarchy as an explicit frame of reference: Educational quality should be promoted first at every school, regardless of racial composition; if bussing aids it, fine; if not, don't waste money on bussing. Participants at another school might reach consensus on this value hierarchy: First priority goes to insuring a second-to-none education for blacks; then to promoting closer association of white and black pupils. The wording is different in each case and reflects different orientations to the problem, but the prospects of merging these two into a common frame of reference appear to be excellent.

Empirical researchers could contribute much to understanding the dimensions of this stage and promoting its effective operation by exploring these questions:

1. What techniques are most effective in helping people identify their own assumptions? Are those suggested above equally successful with various kinds of people in various kinds of situations? Does previous research in other contexts suggest additional techniques to try with this problem?
2. What kinds of values and beliefs are assumed in common by various groupings of people who live and work together? Do they share more of some assumptions than do those who are socially distant?
3. When people appear to differ irreconcilably on certain assumptions of fact or value, what are the most effective ways of dissolving the rigidities? How successful are the following techniques under specified conditions?
    a. Dialoguing in open forum.
    b. Appealing to majority vote.
    c. Arranging for empirical tests of all disagreements over facts.
    d. Living and working together to build a body of shared experiences.

DELIBERATING ON MEANS-END RELATIONSHIPS

Now that the problematical situation has been carefully diagnosed and an explicit frame of reference established, the next step is to derive a firm value judgment from the evidence, the varied interests, and the guiding principles agreed upon. The form of the judgment will be, either implicitly or, better, explicitly, a means-end relationship, and it should be the product of systematic deliberation. Neither the end nor the means is set in advance, but each is created and refined, through imaginative thought, in reciprocal relation to the other. This is dialectical thinking and parallels the transactional view of experience. Let us look at it in greater detail and with illustrations.

The process at this stage has two foci, the selection of ends and the selection of means, but each is studied in close relationship to the other. When the focus is on the choice of ends-in-view, the available means and their costs modify and shape the interests that become the projected goals. For example, an interest in improving the education of black pupils lacks character or form until it is seen as the possible product of bussing them away from their neighborhood school, or of changing the traditional curriculum, or of using unprepared teachers in the white schools, or of tripling the number of black teachers, and so on. The goal of a better education for black pupils through bussing, racially proportionate schools, and encouraging close association of black and white pupils is significantly different from the goal of a better education for black pupils through hiring all black teachers and administrators for predominantly black schools, providing intensive in-service training, and formulating a curriculum to develop pride in being black.

On the other hand, when the focus is on the choice of appropriate means, each possible means needs to be examined for its probable outcomes *beyond the initial end-in-view.* These further outcomes should be viewed as candidates for additional intended goals or as liabilities against the selection of this particular means. For example, if bussing children to racially proportionate schools seems likely to improve the quality of education of black pupils but is also seen as using money that might be spent to purchase needed school equipment, getting parents concerned about the quality of education throughout the district instead of just in a neighborhood school, forecasting a need for the three specialized high schools in the district to become comprehensive schools, and promoting social dating between white and black students, the appropriateness of bussing as a means now depends on whether these other likely outcomes are accepted as additional ends-in-view or viewed as detractions from the worth of the means. In either event, these are the conditions for a responsible choice of means.

This reciprocal or transactional deliberation on possible goals and possible means continues until the participants agree that their selection represents a condition of stable and inclusive interdependence between certain ends and means. The agreement may fall short of consensus, perhaps consisting of a majority vote or accepting a special committee's recommendation. The selection may also fall short of a complete solution to the problem as originally diagnosed. For example—to continue our illustration—the conclusion agreed upon at this point might be (a) to try out bussing at only a few schools, if the majority of parents involved are willing, or (b) to bus children only to those schools with superior records for effective teaching, or (c) to promote bussing experiences for *all* children through greater use of community resources and interschool projects. In fact, the original problem identified at the diagnosis stage may now be significantly altered as a result of this deliberation on ends and means.

This is the stage where the *making* of a value judgment reaches maturity. It is fulfilled when coherently organized means appear comprehensively united with a coherent set of goals to support a proposal for action. Its testing comes next, but the value judgment amounts to a competent declaration of the worth of trying out this proposal in practice.

The dimensions of the deliberation described in this stage rest at a number of points on empirical questions and hypotheses. Scientific investigations by educational researchers at these points could not only test the empirical assumptions in the theory but also improve the techniques for conducting this transactional kind of deliberation. Some possibilities are suggested in the following questions.

1. To what extent, under what conditions, and for what types of people does the perception of a new but possible means-end relation affect or alter an earlier judgment?
2. Granted that the perception of more of the means involved in reaching a desired goal clarifies the meaning of that goal, does it also heighten the anticipated enjoyment of achieving that goal?
3. What could be some unobstrusive measures of "anticipated enjoyment"?
4. Are people less sure of their affections for a possible goal (i.e., a future state of affairs desired) when they know little about the means required than when the goal is conceived in terms of estimated means to reach it?
5. Are there definite skills involved in estimating the other probable

consequences of means besides the achievement of the ends-in-view? If so, can they be specifically taught? In context only? Separate from context?

6. Can the habit of changing one's goals in the light of inadequate old means but of promising new means be taught specifically and separately? Or learned only as a by-product of actual participation in evaluations with a high contingency factor?

TAKING ACTION ON THE JUDGMENT

The purpose of this stage is to test the validity of the value judgment, but there are two distinct senses in which the validity is tested. One is the matter of factual accuracy: Do the perceived relations of the means-end chosen actually work as anticipated when action is taken on its central proposal? The other sense tests the integrity of the participants: Do they believe enough in the validity of this value judgment to put it to the test of public experience?

When the testing action is of comparatively short duration (as in establishing a new dress code for the school), it can be completed before the consequences are reviewed and appraised. In actions covering a considerable period of time (as in promoting social integration of the minority-majority in school), occasional checkpoints may be needed for reviewing and redirecting the program of action. In any case the tryout should be pursued long enough to allow new evidence and possibly new interests to appear. It should also be remembered that this action stage is needed not only to test the validity of the matured value judgment but also to provide data (including changed interests and new factual evidence) for use in making a better judgment next time.

Empirical studies of the degree to which people act on their beliefs and attitudes are already familiar to educational researchers. The techniques which measured this relationship may be appropriate to testing the commitment of participants to the worth of their consensus. Perhaps the greatest need is for the development of natural ways of taking group action on value judgments that insure that all the participants will experience as directly as possible the actual consequences.

REVIEWING THE EXPERIENCED CONSEQUENCES

There may be a temptation for the participants at this final stage to make a gross and global pronouncement: "This was a good decision" or "On the whole, that was a pretty fair judgment." Without denying the relevance and virtue of a summary interpretation, my position is that it should come *after* a reflective examination of the consequences of the tryout. Some people, of course, claim that they do not need this period of reflective examination, that they have an intuitive knowledge of the total outcome that is different from any summation of the detailed parts. This contention may be a matter of personality style, but I would fear that those who rely on a direct intuition of the whole value of the act may be missing some actual occurrences in the process that would significantly alter their final interpretation. At any rate, the richer prospects for scientific investigation into this concluding stage seem to lie in the behavior of those who conduct a reflective examination of the consequences before they reach a global interpretation.

Certain aspects of the experienced consequences deserve special attention because of their evaluative significance. First, of course, is identifying the particular likings and dislikings, the enjoyments and the hurts which occurred as products and by-products of taking action on the judgment. Let us suppose, for example, that the following outcomes were noted in particular:

1. After nine months of bussing, there was no exceptional gain (or loss) in standardized test scores on basic skills, but there was a small decline in the incidence of juvenile delinquency.
2. Most children enjoyed the bus rides, but the bus drivers complained they found it difficult to control the exuberance.
3. Many parents found they enjoyed the longer period of time that their children were away from home.

Which of these outcomes were expected (thus confirming those parts of the judgment)? Which were unexpected (some serving to refute the judgment and others to expand the intended goals next time)?

Second—and of equal importance—is noting how well the predicted factual relationships worked out. Suppose, for example, that some parents feared social intimacy among black and white pupils, and other parents welcomed it, but all agreed that bussing the children would promote it. Then suppose the results were that, while proportionate racial representation was achieved statistically, segregation-by-choice in friendships, playground games, and social activities obviously persisted in all schools but one, which had initiated a systematic program of integrated social activities. What empirical connections were confirmed? Disconfirmed? Left obscure? Suppose further that some unexpected relationships appeared: (a) Teachers found the pupils more interested in classwork that was related to experiences on the bus trips. (b) There was a surprising increase in the number of teachers who expressed a preference for teaching special classes for underprivileged children, whether black or white. Can the unexpected results be readily attributed to certain means used, or does locating the cause require further investigation?

In sum, a global interpretation is now called for. It could be sought in a summary of the answers found to the questions listed above, but such an interpretation is too likely to be static. It could also be sought in the affective terms now being used by the participants—words of praise, of disappointment, of enthusiasm, or of resentment. But these terms are vague, unstandardized, and poor communicators to others who did not themselves undergo this six-stage process. To get a behavioral clue and to highlight the continuing, progressive nature of all value judgments, possibly the best way to learn the group's global interpretation is to ask: After the participants have undergone the consequences of their value judgment in action and have discussed or reflected on the results, how much do they desire to continue the tryout of this proposed course of action (i.e., continue testing the value judgment)? Or how much do they seek reexaminations and revisions of earlier stages in the process? Or to what extent do they prefer to formulate a new and different judgment for the next tryout?

The most needed empirical research in this concluding stage consists chiefly of methods and techniques. What are dependable techniques for identifying objectively the enjoyments and the hurts, the satisfactions and the dissatisfactions? How can the emergence of new interests during the valuation process be detected by unobtrusive measures? A great part of this reviewing stage, however, consists of straight empirical questions—the connections between means tried and results obtained. These will hopefully attract the interest and participation of educational researchers, who will recognize the contextual significance of these questions being the closing phase of the process of making and testing value judgments.

CHAPTER XI

The Limits of Aesthetic Inquiry: A Guide to Educational Research

DAVID W. ECKER
and
EUGENE F. KAELIN

One of the current dogmas concerning educational research is that only methods employing strictly verifiable quantitative measurements of human behavioral patterns are capable of improving educational practice. Although we have no quarrel with precision, exactness of measurement, and continued attempts at scientific replication, we do have some grave misgivings about measurement for the sake of measurement, i.e., about the ineptitude of measuring without a deep prior concern for the intelligibility of what is to be measured in any careful scientific observation. Controls there must be, and correlations must be drawn; but there must also be a careful prior analysis of any area of relevant human experience if we are to insure that we are measuring what we claim to measure.

Traditionally, conceptual analysis of a kind we shall be suggesting has been performed by philosophers and has constituted in recent years the principle concern of the philosophy of education. This relatively new kind of investigation has already succeeded to a degree in combining the specific competencies of philosophers with those of empirical researchers to the practical end of improving education.

Our aim here, as philosophers of art, is to specify the range of activities that may be relevant in determining a peculiarly aesthetic domain of future educational research. In so doing we shall be attempting a necessary first step toward a broadened concept of educational research, one that is defined materially by criteria of relevance set up by related subject matter disciplines and, consequently, one that would indicate the limits within which any formal method may appropriately be applied, if our restriction of prior intelligibility is to be met.

For purposes of convenience only, we shall begin to clarify our domain by referring to those aesthetic experiences that are controlled by close attention to the observable properties of art works and shall provide in three quotations examples of various kinds of language used by critics to report their appraisals of an aesthetic object. For longer critical or theoretical approaches to art, references are made to standard works.

I

Within the area of aesthetic experiences, nothing has ever been quite so obvious as the differences in aesthetic taste, not even the almost equally obvious and confusing array of aesthetic opinion to which disputants may refer in justification of their judgments. The area is definitely gray, one in which fact and fancy are well-nigh indiscriminable, and skepticism is perhaps the safest way of avoiding error. But where skepticism as an attitude may always be considered healthy, as a conclusion taken in advance of inquiry it is always self-defeating and necessarily false. It merely serves to make a fetish of our own ignorance, closing off, as it does, the only way open to us for gaining the knowledge we seek. This initial doctrinal skepticism is precisely the position of the person who finds himself in that unhappy situation in which he is self-consciously certain of his own personally felt feelings and virtually nothing else. But he can be led out of his quasi-religious solipsism if only he keeps talking in order to convince us or, better, himself of the truth of what he says.

Submerged in this dialogue, not only do we experience objects of aesthetic preference, whether they be natural or art facts, but we also talk about them, describe them and our reactions to them, and attempt to justify our aversion or attraction to them. But language is a dangerous tool; it can be used to describe without evaluating or to evaluate without describing the facts of our experience. And some words, like *good*, *aesthetic*, and *art*, can function both ways at once. As a result, some students of the arts, as well as some productive artists, find it difficult to see anything but redundancy in the expression "good work of art." If something is art, then it is good; and nothing that is not good can be considered "art" in this sense of the term. Yet it takes only a moment's reflection to see the error in this supposition. Not everything produced in the name of art succeeds in creating a context of meaningful experience. We call such objects "bad art" without self-contradiction, however, because we use the term *art* descriptively to point out a field of inquiry. In doing so, we are merely answering affirmatively the question, Is a painting, a poem, a musical composition—good or bad—a work of art? Of course they are, since none of these sorts of things is a fact of nature.

We may begin our inquiry, then, by using these terms in a purely ostensive way. They merely point to the kinds of things our inquiry is to be about. Harold Rosenberg, who is good enough a critic to speak, reports in revealing terms his disagreement with Mary McCarthy concerning his own appraisal of action painting. She contentiously maintained that events such as human gestures do not hang on walls, as do real paintings. His retort is right on:

> In reply, I should like to point out that in dealing with *new* things there is a question that precedes that of good or bad. I refer to the question, "What is it?"—the question of identity.[1]

However, he complicates the issue in the very next sentence when he writes,

> To answer this question in such a way as to distinguish between a real novelty and a fake one *is itself an evaluation*, perhaps the primary one for criticism in this revolutionary epoch when art, ideas, mass movements keep changing their nature, so that their most familiar features are often the most misleading.[2]

In this claim he all but concedes victory to the opposition. For how can one answer the question of identity in such a way as to distinguish a real novelty from a fake one? Only by comparing the logical contender with a real, known novelty of the past. Art historians, of course, glory in such investigations, and not all of them assume that novelty, real or fake, constitutes the basis for an evaluation of the thing so identified.

[1] Harold Rosenberg, *The Tradition of the New*, 2d ed. (New York: Horizon Press, 1960), p. 5.

[2] Ibid., pp. 5-6.

From this point of view, the tradition of the new may furnish the content of a historical inquiry, and statements of both fact and value may have their place therein. But Rosenberg is a critic. Can these roles be clearly delimited? Yes, if once again we continue to talk and our talk is a report on the multiform phenomena investigated in careful aesthetic inquiry. As an alternative, we may find ourselves faced with a single option, epitomized in the sculptor-critic's statement: "If someone says his work is art, it's art."[3] Believe it or else. Like Mary McCarthy, however, we should be forgiven for believing, instead, that art is what an artist has made. In the making, he may also have made a judgment—at least he must have known when to stop—that what he had done to that moment was worthy of our contemplation.

We may, of course, ultimately be led to agree with the artist's implicit or even his stated belief, but the issue will be resolved only on the basis of our own experience of the qualities found in the object he has made. To assume otherwise is to force us into an unfortunate dilemma: either to accept the authority of the artist, critic, historian, or whoever else may make claims to knowledge without warrant, or to fall back upon our own skeptical solipsism, which is equally as dogmatic in that it denies in principle the very possibility of such a warrant.

II

Art objects, descriptions of their identity, evaluations by critics, and historical descriptions of changing aesthetic standards—all have a place in the field of aesthetic inquiry. Our immediate task is (a) to show how it becomes possible to work out of the untenable choice between skeptical solipsism on the one hand and uncritical acceptance of authoritarian pronouncement on the other; (b) to construct if possible a scale or schema of the structures of aesthetic knowledge, founded upon the commonality of communicating worlds of experience; and (c) to indicate how future educational research may be undertaken to incorporate the results of our analysis within the institutional framework of professional education.

---

[3] Don Judd, in *Conceptual Art and Conceptual Aspects: April 10 to August 25, 1970*, catalogue compiled by Donald Karshan (New York: New York Cultural Center, 1970), p. 47.

Consider the first objective. The untenable positions may be thought of as two circles and their possible relations. Where the two circles do not overlap, one of them represents the skeptic's world, in which the limits of his knowledge are exhausted by his own personal, lived reactions either to the individual creative efforts of artists or to the ever changing styles of practicing artists; doctrinal skepticism arbitrarily cuts off one world—that of the skeptic—from all other possible worlds encapsulating the experience of others. One man alone, impervious to all external influences, can do nothing but contemplate the figments of his own imagination; he cannot even be assured that whatever private language he might invent would refer to anything at all.[4]

How reassuring, therefore, is the position of the dogmatist who, on the other hand, claims that absolute knowledge of matters aesthetic is warranted through the application of universal rules obtaining throughout the history of human culture. Concerning the knowledge of the rules themselves, we need only attend to the rhetoric. The dogmatist's world overlaps with all other possible worlds only by fiat; he has his knowledge by infusion much in the same way as earlier epistemologists claimed that the knowledge of first principles was intuited and not proved at all.

But questions invariably arise; and when the question is asked how one warrants any claim to knowledge of universal rules, one must appeal to an ultimately self-evident principle or be thrown back upon the tentative, struggling efforts of finite human minds to know anything at all—proceeding from the observation of particular cases through idealizations to more or less limited generalizations, and from these to predictions of things to come. Nothing, in a word, corresponds in experience to a grand intuition of a universal truth, by which aesthetic knowledge is infused into all subjects of equal sensitivity. It is for this reason, no doubt, that a dogmatic critic casts doubt upon the sensitivity of his own skeptical readers, those whose discomfiture forces an inquiring attitude into the dogmatist's suppressed questions. The empirical researcher, in particular, will want to make correlations between a standard set of observations based upon the reactions of an aesthetically knowledgeable group and another set based upon those of a test group. What effect does time, variance of past experiences, differences in acculturation, socioeconomic status within the society, not to mention the more fleeting of parameters—psychological sets and attitudes, environmental conditions, indeed, the physical state of the organism—have on a subject's ability to report correctly on his own situation. And to what degree are we ourselves capable of judging the reports of others so that a meaningful correlation may be drawn? Obviously, an untold amount of inquiry into the conditions under which aesthetic objects are created and appreciated, as well as those under which this communicative process functions within the nexus of the other institutions of society, must be performed in order to reassure ourselves that no gray areas of uninterpreted experience remain in either of the two communicating worlds, whose degree of communication is established precisely by the degree to which the empirical researcher has succeeded in establishing a correlation between the responses of his standard and his test groups.

---

[4] Cf. Ludwig Wittgenstein, *Philosophical Investigations*, trans. G. E. M. Anscombe (New York: Macmillan Co., 1953), secs. 243 ff.

In other words, only by performing aesthetic inquiry, by asking and answering the heretofore unanswered questions, can we be satisfied that the world of our aesthetic experiences is as we claim it to be. The predicament of the dogmatist turns out, upon analysis, to be no less grounded in personal experience than that of the solipsist who claims he can know nothing else. Taken as a statement of the ideal concerning aesthetic knowledge, there should be little argument; taken as a proposal of method, the kindest thing we may say about it is that it is unrealistic indeed.

In sum, should we choose to be realistic, we must admit what we all are convinced is the case: Two worlds of human experience communicate because of shared contents and structures, and the purpose of inquiry is to broaden the scope of this communication. Even the artist's reaction to his own work need be no different from that of any other viewer who knows what he is looking at. What makes it possible for communication to take place at all is the existence of a shareable, public object whose very structures control all relevant responses to it.

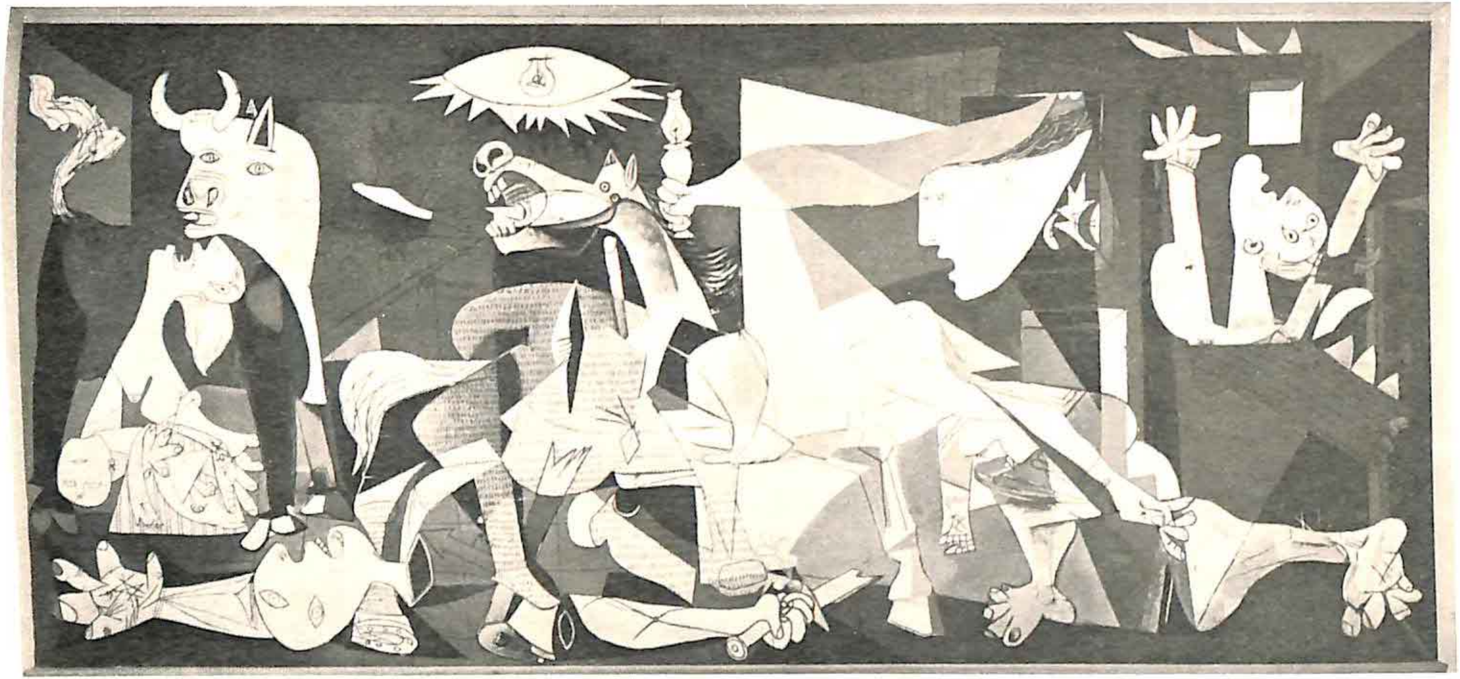
Let us say we are standing in the Museum of Modern Art before Picasso's *Guernica*. See Figure 1. There is a primary aesthetic object. If another person can follow our indication of it, we have established communication. Moreover, if we can believe their reports, scientific investigators of whatever stripe have nothing more than this sort of common relation to a single world from which to draw their own theoretical explanations. For this reason all language used to describe the public objects of human experience must be grounded in ostensive definitions—those which point to their objects. However elaborate the linguistic system scientists may develop, they can have no more assurance of its applicability to fact than that accorded to its disclosure of primary referential objects. In this sense a scientific theory is an indicator of experiences to come. As David Hume put it, the practicing scientist merely bets that the future will resemble the past.

Within aesthetic inquiry we may readily admit that an artist creates the primary objects of initial reference in his initial gesture of communication, and it follows thereupon that the first job of the critic who would give evaluation of any kind is to pose and answer the question of an art work's identity. Our assumption as aesthetic inquirers, then, is as follows: All aesthetic knowledge must derive from a correct description of such basic art facts. Whether or not the gesture of the artist and our responses to his product may be meaningfully said to constitute a language,[5] it is bootless to deny that we must use language, either as already constituted (as ordinary language philosophers insist) or construct a more restricted, that is, quasi-technical, terminology (as other philosophers and virtually all scientists do) to communicate to another person what our reaction has been and what, presumably, his ought to be.[6] Disagreements concerning ultimate evaluations of aesthetic objects must therefore be referred to a system of possible agreements on the descriptive adequacy of our aesthetic language. It is only in this way that the areas of overlap between communicating worlds of descriptive knowledge can be guaranteed as *knowledge* and extended ultimately to include all those worlds of the common-

---

[5] See R. G. Collingwood, *The Principles of Art*, paperback reprint (New York: Oxford University Press, 1958).

[6] See E. F. Kaelin, "Isness and Oughtness," *Art Education* 21 (January 1968): 6-10.



PICASSO, Pablo. *Guernica.* (1937, May-early June). Oil on canvas, 11′ 5½″ x 25′ 5¾″. On extended loan to The Museum of Modern Art, New York, from the artist.

FIG. 1.—Picasso's *Guernica*

language users. Here, as always, knowledge by description depends upon a prior knowledge by acquaintance.

Although the extension of overlapping meaningful worlds may be treated anthropologically as a culture tied to a specific place in time, the danger of elevating common attitudes and shared meanings of critical discourse into an arbitrarily restrictive elitist culture should be apparent and the temptation forcefully resisted. For, if a critic must use language to make his meanings clear, any person capable of understanding his language may, by consulting his own experience, just as readily focus upon those features of the public object referred to in the critic's use of language. Thus, the only ground we would have exposed to warrant taking any position vis-à-vis the current elitist-populist controversy of the New York art world is the following: Elitists make claims to possessing the insights required of warranted aesthetic knowledge, as of course they may; and populists rightly insist that these insights may and therefore ought to be shared.[7] And they must indeed be shared, if we are to demonstrate the validity of our claim to the universalizability of aesthetic knowledge. For starters we need only begin with descriptions of the objects artists have made.

If the polarity of human attitudes expressed in skeptical and dogmatic epistemologies describes the outer bounds of aesthetic inquiry, it seems clear that the basic limit—found as that condition allowing communication between the various worlds of aesthetic experients—is to be located within the responses of these various subjects to a single aesthetic object. From this starting point we move to the second phase of our project, which is to show how, once a unique "object of criticism"[8] has been isolated for aesthetic contemplation, the use of language to describe or evaluate the art work affects what can be known through aesthetic inquiry.

In order to present a synoptic view at the start, we shall arrange the orders or levels of discourse that may be associated with any knowledge claims concerning the existence or interpretation of aesthetic objects. A quick glance at Figure 2 below will reveal that each higher level of discourse has as its object of reference either one or more, perhaps all, of the strata located beneath it. This stratification will be useful to outline the limits of the various roles or professional attitudes, various techniques and methods proper to each of these, and finally the routing necessitated for checking any truth claims made at any one level. The orders of knowledge discriminable within the diagram will vary depending upon our point of departure beginning, for example, from either the top or the bottom of the "ladder." From any superordinate position, the warranted knowledge claims of a subordinate stratum will be taken as implicit, whereas the implicit knowledge of the subordinate strata is raised to the level of explicitness at that level of discourse. This means, in effect, that if too many rungs of the ladder are missing and if one's aim is to achieve explicit knowledge of aesthetic objects, then "one just can't get there from here."

---

[7] Quoted from remarks by Hilton Kramer, Barbara Rose, and Lawrence Alloway at the Conference on Art Criticism and Art Education held under the auspices of the Department of Art Education, New York University, at the Guggenheim Museum, May 15-16, 1970. Proceedings published under title, *Art Criticism and Art Education*, forthcoming.

[8] See Stephen C. Pepper, *The Basis of Criticism in the Arts* (Cambridge, Mass.: Harvard University Press, 1956), pp. 168-69.

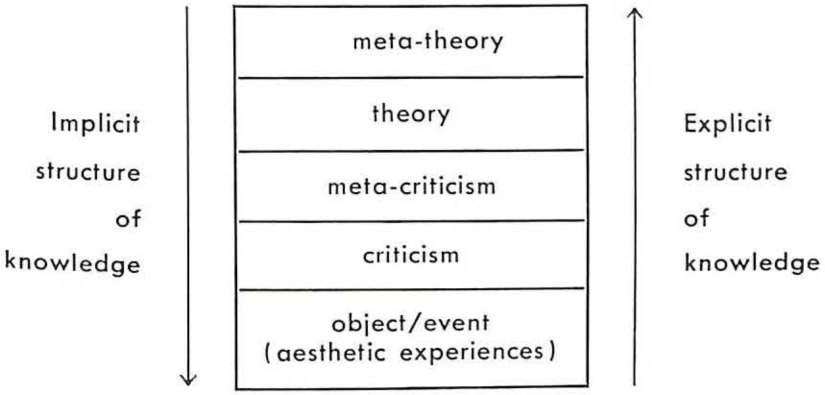
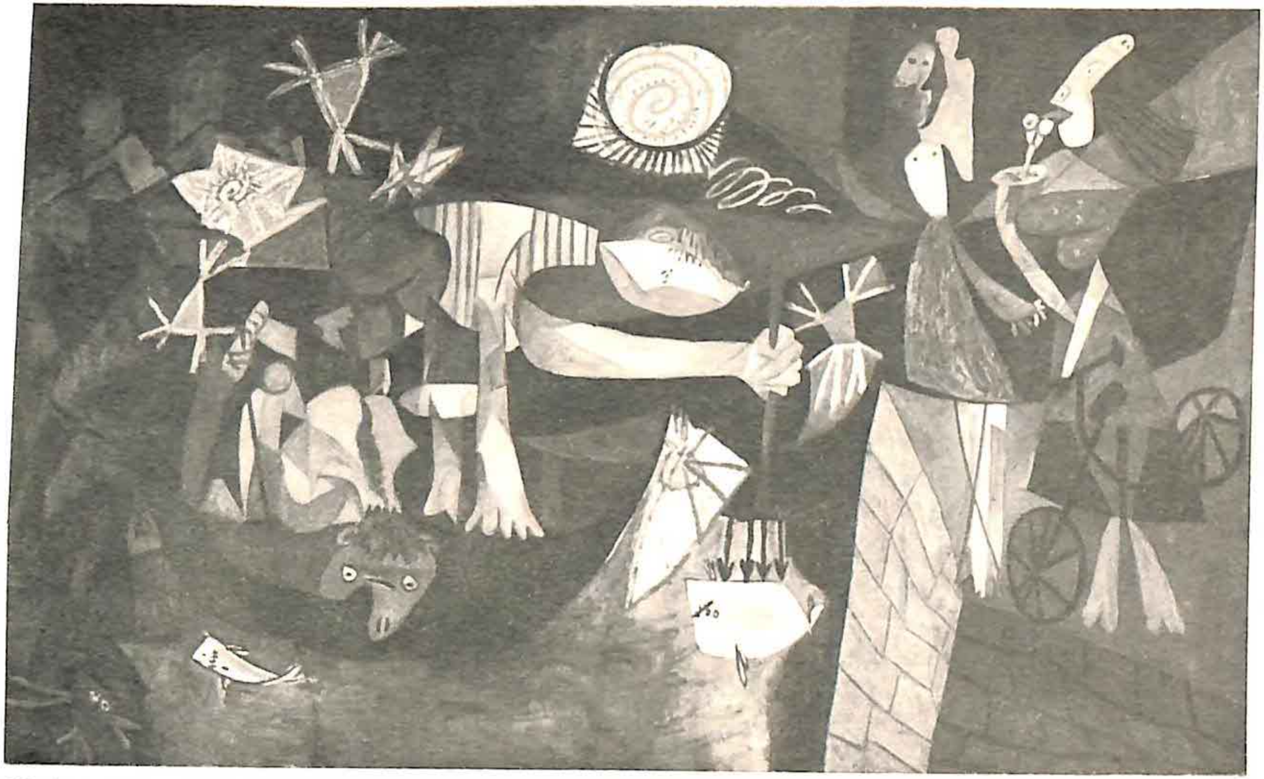

Implicit structure of knowledge
meta-theory
theory
meta-criticism
criticism
object/event (aesthetic experiences)
Explicit structure of knowledge

FIG. 2.—Levels of discourse

To begin at the bottom, aesthetic inquiry must be grounded, as any other would-be scientific investigation, in the data of observation. But it matters greatly which data. Ultimately, all theorists who would also be practitioners must be able to point out the facts referred to in their discourse, or, in Peirce's terms, to devise a system of indexical terms to "tie theory to nature." The facts of aesthetic experience, qua aesthetic, must therefore constitute the data for aesthetic theory. In their attempts to meet this primary



PICASSO, Pablo. *Night Fishing at Antibes*. (1939, August). Oil on canvas, 6′ 9″ x 11′ 4″. Collection, The Museum of Modern Art, New York. Mrs. Simon Guggenheim Fund.

FIG. 3.—Picasso's *Night Fishing at Antibes*

condition of referential adequacy,[9] some aestheticians have gone beyond the facts of actual response and substituted the data of other disciplines as explanations for the existence of aesthetic objects. Cases in point are George Santayana [10] and, in some of his unwary moments, D. W. Prall.[11]

Naturalistic explanations of aesthetic objects, however, are reductionist in that they attempt to explain the occurrence of aesthetic data in terms of natural—read "physical," "physiological," "biological," or, most often, "psychological"—phenomena. Thus, color is associated with light waves and their various physical properties, tones with sound waves and their properties, and so forth. Since such physical properties do not enter into the psychologically unique response centers of stimulated organisms, quite obviously this discourse is located at the level of theory in our scheme, and is in fact notoriously unrewarding for any inquiry occurring at a lower stratum of activity, notably that of criticism, in which critics merely give reports on the qualities they have experienced. Moreover, since it gives them no indication of how to look at the phenomenal object before them, critics have some ground for complaining against this sort of empty-headed empiricism.

For an example of physiological theory which preempts the place of straightforward critical response, consider the statement of George Santayana, who writes:

> The eye, by an instinctive movement, turns so as to bring every impression upon that point of the retina, near its centre, which has the acutest sensibility. A series of muscular sensations therefore always follows upon the conspicuous excitement of any outlying point. The object, as the eye brings it to the centre of vision, excites a series of points upon the retina; and the local sign, or peculiar quality of sensation, proper to each of these spots, is associated with that series of muscular feelings involved in turning the eyes. These feelings henceforth revive together; it is enough that a point in the periphery of the retina should receive a ray for the mind to feel, together with that impression, the suggestion

---

[9] David W. Ecker, "Justifying Aesthetic Judgments," *Art Education* 20 (May 1967): 5-8.

[10] George Santayana. *The Sense of Beauty* (New York: Modern Library, 1955).

[11] D. W. Prall, *Aesthetic Judgment* (New York: Thomas Y. Crowell Co., 1929), and *Aesthetic Analysis* (New York: Thomas Y. Crowell Co., 1936).

of a motion, and of the line of points that lies between the excited point and the centre of vision.[12]

Clearly, this is physiological theory with the slimmest of ties to the phenomenal character of aesthetic experiences. Moreover, should any practicing critic be seduced by the blandishments of these naturalistic abstractions, we may be convinced that he has been deluded by a true, but unfortunately irrelevant, knowledge claim. Knowledge of the function of the rods and cones of our visual apparatus is certainly available; and, moreover, we can be assured that if they do not function, we will see nothing. But what it is that we have seen once they have been stimulated goes unelucidated in the naturalistic explanation.

On the other hand, critics, who sometimes are known to have used their eyes, use ordinary language to describe what they have seen; but even they may be deluded by a doctrine of historical development of styles or modes of expression and couch their explanations in the theoretical terms of historical analysis. We may call this the art-history fallacy of aesthetic explanation. Compare the experience of Picasso's *Night Fishing at Antibes* (Figure 3) with the historian Levitine's account of it (Excerpt 1):

> In spite of the distortions and abstractions, the artist's approach cannot be explained away as a mere "play of form." Thus, since *Night Fishing* contains a recognizable subject, as well as an intriguing manipulation of form, and is—at least from this point of view—a painting *comme les autres*, one may perhaps be justified in submitting it to the trials of a study of filiation—a typical artifice of art historians.
>
> Picasso's theme shows a degree of kinship with a variety of well-known "fishing subjects," ranging from Raphael's *The Miraculous Draught of Fishes* (Victoria and Albert Museum) to Copley's *Watson and the Shark* (Boston Museum of Fine Arts). However, this kinship is superficial. In a less obvious, but far more significant way, Picasso's composition appears to be more closely related to a lesser known work of an earlier tradition: *The Bathers*, a seventeenth century anonymous Dutch painting in the Louvre.

EXCERPT No. 1.—From George Levitine, "The Filiation of Picasso's *Night Fishing at Antibes*," *Journal of Aesthetics and Art Criticism* 22 (Winter 1963): 172

---

[12] Santayana, *Sense of Beauty*, pp. 90-91.

Just as obviously, as in the case of an appeal to physiology, this move is a snare and a trap enticing the critic to switch roles with the historian, whose specific mode of inquiry is located properly at the level of theory. In short, aesthetics as traditionally practiced, science, and history—all have a legitimate interest in aesthetic objects, but none has been a successful substitute for perspicuous criticism or authentic appreciation. One simply cannot bypass the stage of defensible art criticism.

Good critics speak in ordinary language, and when they communicate they call our attention to the phenomenal properties of the natural object. This is the process by which two centers or worlds of human experience are, as we have said above, to communicate. And the degree to which this communication takes place is a measure of the expansion of the shared contents of experience as each partner in the communication moves from the state of skepticism to ultimate self-conscious assurance in the contents of his knowledge. The only restriction we should like to place on this process is that the critic's audience test each critical assertion against the appearance in his own experience of the designated phenomenal properties. This is the task, along with some metacritical reflections, assumed by the late Douglas N. Morgan in his criticism of Picasso's *Antibes*, in Excerpt 2:

> Picasso never wholly deserts our world. He plays upon and against our historic expectation-patterns. But the people he makes—while surely people and not mere symbols—are *pictorial* people, people to be seen and not sympathized with. As pictorial, they are not really distorted at all. They are exactly the kind of people they were created to be. They don't look "wrong"—they look "right." They are not "misshapen"; they are correctly shaped, in their pictorial context. These boys and girls have no antecedent history against which to speak of "distortion." They are what we see them to be, no more and no less.

EXCERPT No. 2.—From Douglas N. Morgan, "Picasso's People: A Lesson in Making Sense," *Journal of Aesthetics and Art Criticism* 22 (Winter 1963): 169

In a similar vein, it is one thing to mention that the *Guernica* is a mural-sized painting (11′6″ × 25′8″), that its creator was Picasso, even that it was done in only one dimension of color-value (black, white, grays), still another to perceive the frightening explosiveness of the jagged forms and the way in which they express the destructiveness and consequent horror of modern warfare. (See Figure 1.) The language of the first description is physicalistic and value neutral; of the second, expressivistic and necessarily value laden.

The distinction drawn here, between the natural object as it has been formed by the artist and the created "object" that is a synthesis of its phenomenal properties, has likewise been the source of some confusion. Recall the argument between Mary McCarthy and Harold Rosenberg. Her insistence that only paintings hang on walls, and not human gestures, indicates to us that she was referring to the physical object manipulated by carpenters (in crating it), by movers (in shipping it), and by curators (in hanging it). Rosenberg, on the other hand, insisted that the phenomenal properties of action paintings viewed by him disclose a pattern of human gestures made in forming the object. The confusion laid out in these terms shows how there was no argument taking place between the two concerning the nature of criticism, but only about what in theory is to count as an aesthetic object. From Rosenberg's point of view, the task of the critic is to look upon the work as an action perpetrated by the artist:

> At a certain moment the canvas began to appear to one American painter after another as an arena in which to act—rather than as a space in which to reproduce, re-design, analyze or "express" an object, actual or imagined. What was to go on the canvas was not a picture but an event.
>
> The painter no longer approached his easel with an image in his mind; he went up to it with material in front of him. The image would be the result of this encounter.[13]

This description of the working stance of practicing artists, then, becomes a prescription for the way in which it is proper to view the work so produced. As such, the description is critical; the prescription, metacritical.

Anyone who takes the confused interchange between Rosenberg and McCarthy as an argument may be said to have committed the fallacy of discursive displacement, which English-speaking philosophers refer to rather dully as "a category mistake." Attending to one's discourse in order to avoid the category mistakes of discursive displacement immediately places one at the higher level of discursive usage. If the mistake occurs in criticism, the corrective stance is always to be found in metacriticism; if in theory, then at the level of metatheory. The only assurance, however, that metacriticism or metatheory, whose one-distance removal from a primary discursive level allows it to describe the functions of language at that level, will allow us to correct any mistakes made there, derives from the logical properties already implicit within the primary discursive levels; and the primary discursive levels themselves have significance only in the degree to which their assertions exhibit empirical adequacy. Thus criticism and theory must be both true and relevant.

---

[13] Harold Rosenberg, *American Painting Today* (New York: Horizon Press, 1965), p. 25.

In other words, there is no infinite regress in the explanation of aesthetic explanations because criticism and theory are grounded in fact. Metacriticism and metatheory are themselves grounded in the logical structures of the language used in criticism and theory. In this way all discourse having any relevance to aesthetic inquiry is grounded either in fact or in the uses of language to refer to fact. The present inquiry, constituted by a use of language concerning all uses of language relevant to aesthetic fact, may be viewed—indeed must be—as an empirical investigation of two differing but related sets of facts: first, the phenomenal appearance of aesthetic objects, and, second, the language or languages that may be used to account for their appearance.

If we are right in our interpretation of the expression, the structures of language levels related to the primary experiences of aesthetic objects constitute the relevance of the "structure of knowledge" (or its theory) for continued educational research within the domain of aesthetics proper. Given the fact that we have located at this point five levels of significant institutionalizable activity, ranging from the creation and appreciation of art to criticism or theory and from there to metacriticism or metatheory—all of which return us to our basic experience of these facts as the ultimate warrant for any knowledge claim concerning them—there would seem to be nothing more that might significantly be said on the subject. And if this is true, as we hope it is, educational researchers need not look beyond this schema to conceive an area of research but must look, instead, within it.

IV

It will be our purpose here to indicate the methodological considerations necessary to delineate the procedures by which the aesthetician is motivated to move up the explanatory ladder. We may begin with an account of "the aesthetic object," the experience or creation of which constitutes our first step.

One of the most illuminating pieces of aesthetic writing is Paul Ziff's " 'Art' and the Object of Art." [14] He insists, rightly, that the traditional distinction between the natural and the so-called aesthetic object rests upon a metaphysical confusion. Metaphysics, of course, is another theoretical approach to aesthetic analysis. Introducing metaphysics into our scheme, however, only confounds the issue at hand. Realists argue with idealists in much the same way that McCarthy "argued" with Rosenberg. If we use the term *object* to refer to the ontological entities of the real world, common sense will show that there are not two objects, but only one. What the realist does is to give a different description of this object, one and the same as that which appears within the idealist's experience. In recent aesthetics Susanne Langer has made much of the distinction to elucidate the difference between the material conditions of creation and the appearance of "virtual" aesthetic objects in a variety of modalities.[15] If, however, we pay close attention to our own uses of language, it becomes very easy to distinguish, even within ordinary usage, between a description of a real object and, for want of a better term, the phenomenological "object of criticism."

Phenomenology, of course, has never made a claim to expositing a metaphysical truth; it purports only to provide a method by which anyone may be led to isolate the relevant characteristics of the phenomenal art object. Once the phenomenologist has suspended belief in the naturalistic "explanations" of his experience and concentrated upon describing the characteristics of that experience, he like anyone else must use ordinary language to describe what appears. In order to perform this feat he must learn to "bracket out" all the irrelevant knowledge at his disposal, whether it be the beliefs of common sense or any logical or scientific or historical knowledge he already possesses. In other words, he merely leaves himself open to the experience of the object in its phenomenal appearance. The first movement from step one of our ladder, therefore, is permitted to anyone practicing this so-called phenomenological *epoché*.[16]

---

[14] In William Elton, ed., *Essays in Aesthetics and Language* (Oxford: Blackwell, 1954), pp. 170-86.

[15] See her *Feeling and Form* (New York: Charles Scribner's Sons, 1953).

> The examination takes the form of a descriptive inventory. We cling as closely as possible to what is directly given to the eyes, on the assumption that in a successful painting the essential meaning is directly expressed in the properties of the visual form. Picasso's painting is composed of three main areas. The narrow panel on the left presents the town and medieval castle of Antibes and continues as a vertical string of rocks which frame the water of the harbor. There is the central medallion of the two fishermen in their boat, surrounded by lights and fishes. And the panel on the right shows two girls on a stone jetty. In the projective frontal plane of the canvas the fishermen hold the center of an approximately symmetrical composition. . . .
>
> . . . .
>
> the theme of the fishermen carries considerable weight not only because of its size and central position but also by its almost architectonic stability, which makes the busy men in the floating boat look, paradoxically, more solid than the stone constructions to their left and right. . . .
>
> . . . .
>
> Whereas the facial features of the upright man are narrowed in restrictive concentration, the leaning man's eyes are open and far apart, and his nose reaches probingly down into the water, which looks empty but may hide anything. The pregnancy of water is familiar to the student of dreams, who speaks of water as a maternal symbol. In comparison with the three-dimensionality of the fisherman helplessly scrutinizing the future, everything else loses reality. The spearman is stopped in a frozen gesture, the stones of the historical castle

---

[16] See E. F. Kaelin, "Epoché and Relevance in Aesthetic Discourse," in *Aesthetic Concepts and Education*, ed. R. A. Smith (Urbana, Ill.: University of Illinois Press, 1970), pp. 135-63.

> recede, the girls flatten out. The foreboding of violent, but unknown, things to come emerges as the dominant theme of the painting.
> (Reprinted in Rudolf Arnheim, *Toward A Psychology of Art: Collected Essays* [Berkeley: University of California Press, 1966].)

EXCERPT No. 3.—From Rudolf Arnheim, "Picasso's *Night Fishing at Antibes*," *Journal of Aesthetics and Art Criticism* 22 (Winter 1963): 165-67

As in Excerpt 3, the critic's other foot is placed securely on the second rung of the ladder when he uses language in such a way as to identify the objects of his attention. And it is not uncommon for him to clarify his own experience at the same time he clarifies that of others—by finding the words to describe what has appeared to him. This is what Rosenberg should have said in reply to Miss McCarthy; for his reference to action painting, as did Bernard Berenson's reference to "tactile values,"[17] succeeded to an eminent degree in calling attention to the structures of two different kinds of aesthetic experience. In this respect we may say of aesthetic objects, as Ludwig Wittgenstein said of the real world,[18] that the universe is what is the case and, by implication, that aesthetic objects or the world are, linguistically, the sum of all the statements that may be correctly uttered concerning them. This is not to claim, of course, that all possibly true statements about the world or aesthetic objects are actually uttered by the scientists or art critics—only that such objects are known or pointed out by the open set of descriptive statements properly predicated upon them.

Metacriticism enters the scene when critics differ and attempt to justify their own statements on some basis other than their observation of the facts. Some, for example, may appeal to a theory of art—historical, aesthetic, or scientific. Others, such as Stephen Pepper, have said: "It follows that good criticism is, as stated at the outset, criticism based on a good philosophy. For a good philosophy is simply the best disposition of all evidence available."[19] But such appeals which are not at the same time reinforced through a reinterpretation of the primary aesthetic experience are to that extent vacuous. If they are not vacuous, the critic's first foot has moved perforce one rung higher on the ladder.

---

[17] Bernard Berenson, *Aesthetics and History* (Garden City, N.Y.: Doubleday Anchor Books, n.d.), pp. 69-73, 84-86, 265.

[18] In *Tractatus-Logico-Philosophicus* (New York: Harcourt, Brace & Co., 1922), *passim.*

[19] Pepper, *Basis of Criticism in the Arts*, p. 15.

One of the common moves of the metacritic[20] is to classify the ways in which criticism has been performed. Most of them are prone to distinguish between levels of critical activity, such as description, interpretation, and evaluation. This conventional differentiation of the levels of critical activity has a limited utility; for what is intended by description is a verbalization of the structures of an aesthetic surface. Unfortunately, however, there are many works of art whose aesthetic surfaces are arranged in such a way as to not only present an abstract sensuous design but also to represent natural objects or even ideas, both iconologically and iconographically. Phenomenologists who persevere in their descriptive tendency refer to these phenomena as a "thickening of the surface."[21] In this view the critic is not so much interpreting (a misleading expression)[22] as he is describing the deeper structures of an aesthetic experience, which, as had, constitutes step one of the ladder. As all the discriminable elements, both of a surface and a depth nature, tend to fund within the context of the experience, the aesthetic object comes to closure. Indeed, some phenomenologists have gone so far as to say that our evaluative terms are merely ways of indicating the degree to which significant closure has been achieved in the experience of the critic. Rather than disputing one's tastes or likes when offering contradictory evaluative judgments, however, critics are describing a complex state of affairs, and one or the other or both may be wrong. And in this phenomenological appraisal of primary critical behavior we find that phenomenology is likewise a metacritical discipline.

In using such terms as *good, forceful, clear, powerful, expres-*

---

[20] See Edmund B. Feldman, "Interpretation and Explanation in the Teaching of Art Criticism," in *Art Criticism and Art Education,* forthcoming (note 7).

[21] See E. F. Kaelin, "The Visibility of Things Seen," in *Invitation to Phenomenology,* ed. J. M. Edie (Chicago: Quadrangle Books, Inc., 1965), pp. 30-58, and G. W. Linden, *Reflections on the Screen* (New York: Wadsworth Publishing Co., 1969).

[22] Cf. Susan Sontag, *Against Interpretation* (New York: Farrar, Straus, & Giroux, 1966), pp. 3-14.

sive, and the like,[23] the critic is said to be describing the way in which the object made by the artist controls his experience. The success value implied in the use of positive aesthetic evaluations, in other words, is likewise an objective (read "intersubjective") feature of aesthetic objects, as long as it is the object that controls the experience. And it will, if the percipient continues to attend to the phenomenal character of that experience.

The theoretical justification for phenomenological metacriticism consists in methodology; it prescribes a set of critical procedures and will ultimately be accepted or rejected on pragmatic grounds. Either it will work in clarifying our experience and thus constitute the possible corrective to the misuse of the language of criticism and be accepted on those terms, or it will not and be rejected on the same terms. Alternative approaches have been more "scientifically" oriented, since they have not perceived the abstract character of the so-called natural standpoint,[24] which is precisely the approach bracketed out of relevance by our own initial standpoint. Whatever theories of art may be constructed must be erected on phenomenologically "purified" data. Traditionally, such theories come down to us in a formula, reading for the most part "art as . . ."

Numerous are the historical accounts of art considered as emotion, form, expression, or what have you; and as Morris Weitz points out in his classical metatheoretical piece, "The Role of Theory in Aesthetics," [25] each of these theories may be said to refer to some generalized characteristics of aesthetic experience without constituting either distributively or collectively a set of necessary and sufficient conditions for defining the concept of art. His skeptical conclusion that there can be no theory of art is itself metatheoretical and is grounded in an examination of the logical properties of linguistic explanations; it is subject to the same tests of validity applied to any explanation given at this level of abstraction.

The force of Weitz's conclusion stems from his assumption that it is the aim of theory to produce "real" definitions in terms of a set of necessary and sufficient conditions for the identity of a meaning of a concept. But if this requirement is not placed upon a lower-level theory, other metatheories become possible. Weitz's own *Hamlet and the Philosophy of Literary Criticism*,[26] for example, tends to belie the strongest interpretation of his own metatheoretical thesis, for it is itself a metacritical interpretation of selected historical criticisms of Shakespeare's play. It constitutes, therefore, one coherent exercise of an aesthetic theory, to wit, that various "interpretations" of a single aesthetic object may produce a more meaningful interpretation of the given object. Using the term in its weaker set of implications, this, too, is unavoidably a "theory" of aesthetic objects, metacriticism in practice, and first-order criticism by remote control. We do not fault him for any of these activities, but only for not paying closer attention to his own occasionally brilliant uses of language. The very title of his *Hamlet* piece contains reference to three levels of aesthetic inquiry: the work itself, criticism, and philosophy as a theory of criticism.

---

[23] Cf. Frank Sibley, "Aesthetic Concepts," *Philosophical Review* 68 (1959): 421-50.

[24] See Edmund Husserl, *ideas*, trans. W. R. Boyce Gibson (New York: Macmillan Co., 1931), pp. 101-11.

[25] In *Journal of Aesthetics and Art Criticism* 15 (September 1956): 27-35.

A recent approach to metatheory that is dogmatic where Weitz is skeptical also bears an intriguing title, with an equally interesting subtitle. The title of Nelson Goodman's book, *Languages of Art*,[27] seems at least to promise an investigation of aesthetic discourse. Instead of discourse on the arts, however, it presents the arts, both "autographic" and "allographic," in terms of their susceptibility to constituting a symbolic notation. But notationality is strictly defined in reference to the logical properties of symbolic systems. Thus the author appears to have been more faithful to the aim expressed in his subtitle, *An Approach to a Theory of Symbols*.

Since the ultimate evidence of this study is taken to be symbological, its immediate reference is to symbolic systems, only one of which is taken to be art. With respect to our scheme, therefore, Goodman's work is metatheoretical, i.e., concerned with symbolic systems in general, and will be applicable to art only if works of art are correctly interpreted in the first instance as symbolic vehicles. This theoretical claim, however, is only an assumption of the book. As a result, little can be found of direct relevance to either criticism

---

26. Cleveland and New York: World Publishing Co., 1966.

27. Indianapolis: Bobbs-Merrill Co., 1968.

or metacriticism, not to mention a consummatory aesthetic experience. That is why we read, while looking for a clear description of the characteristics of aesthetic experience, "Three symptoms of the aesthetic may be syntactic density, semantic density, and syntactic repleteness."[28] All these properties of a symbolic system are given careful definition, to be sure; and there is a fourth and final symptom of the aesthetic, called *exemplificationality*, which is likewise given explicit definition.[29] But, as the author admits, these four symptoms are considered severally neither sufficient nor necessary for aesthetic experience; they are, rather, "conjunctively sufficient" and "disjunctively necessary." And he interprets for us, ". . . an experience is aesthetic if it has all these attributes and only if it has at least one of them"[30]—a theoretical conclusion arrived at on the basis of metatheoretical considerations.

Unlike Weitz, for whom the reverse is true, Goodman is aware that the use of his terms, so devised, may not correspond to ordinary usage;[31] but he has produced a precise, logical determination of the term *aesthetic* that has meaning within a system of knowledge concerning symbolic systems. The question remaining, and unasked by Goodman, is whether this sort of metatheoretical precision illuminates or obscures what we do or undergo when having a first-order aesthetic experience. Since he considers the arts as a kind of symbolic system to begin with, they are "explained" in terms of syntactic and semantic properties of symbols.

In effect, Goodman begins his inquiry at the top of the ladder, but fails to bring us back down below the level of theory itself. And this is what might have been expected, given the tactical approach explained in the subtitle to his book. We shall not go so far as to say, along with the wag, that a logician cannot do aesthetics; we are merely stating that metatheory, whether it is an ordinary language linguistics or a more formally logical one, is the last of the relevant strata of aesthetic inquiry, one in which a great deal more research must be undertaken in order to clarify the *contents*

[28] Ibid., p. 252.

[29] Ibid., p. 253.

[30] Ibid., p. 254.

[31] Ibid.

of our aesthetic experience, but it is hardly a fitting starting point for aesthetic inquiry. Quite obviously there are as many metatheoretical approaches to aesthetic inquiry as there are ways to approach the arts.

Perhaps the American aesthetician who has done most to clarify the various strata of aesthetic inquiry is Stephen C. Pepper.[32] Beginning with his definition of a strictly empirical approach, he too attempts to steer a middle course between skepticism and dogmatism, and has written entire books with relevance to virtually every stratum of possible aesthetic inquiry. For example, his *Aesthetic Quality* [33] is offered as an explanation of the direct phenomena of qualitative aesthetic experience (step 1), and his *Principles of Art Appreciation* [34] attempts to clarify the basis for our critical appreciation of works of art (step 2). Next, at the level of metacriticism (step 3), he offered *The Basis of Criticism in the Arts.*[35] Perhaps the title should have read "The Bases of Criticism in the Arts" since he was applying his theory of the four world hypotheses explained more fully in a book called *World Hypotheses, A Study in Evidence.*[36] In this way he has shown how a metaphysical theory or theories (step 4) may constitute a guide for metacritical evaluation.

Pepper's own metatheoretical efforts have always been characteristically empirical. He has attempted to show how the four theories of criticism—formism, organicism, mechanism, and contextualism—operate together to allow a more coherent definition of "the aesthetic work of art," which, as the object of criticism, is the control object for our first-order aesthetic experiences in the first place. This theoretical approach to criticism then culminated in his *The Work of Art* [37] (step 5). What allows Pepper to climb down our ladder, i.e., allows him to tie metatheory to primary aesthetic experiences, is precisely his definition of the aesthetic object as the "object of criticism" which appears in successive manifestations to a subject open to the aesthetic experience. If we are right, all theoretical or metatheoretical investigations of aesthetic experience must "pay off" by returning the investigator to the original facts which initially prompted his inquiry.

---

[32] Cf. his recent statement in the *Journal of Aesthetics and Art Criticism,* 28 (spring 1970): 275-86.

[33] New York: Charles Scribner's Sons, 1937.

[34] New York: Harcourt, Brace & World, 1949.

[35] Cambridge, Mass.: Harvard University Press, 1945.

[36] Berkeley and Los Angeles: University of California Press, 1942.

[37] Bloomington, Ind.: Indiana University Press, 1955.

Thus, there seem to be two generic ways of failing in aesthetic inquiry: the most basic is noted in the apparent irrelevance of an investigation not prompted by the facts obtainable in a first-order aesthetic experience; the second is not to return thereto. It is not being claimed here that any one investigator need go the whole route; as a matter of fact, most contemporary aestheticians write narrowly conceived essays on specific points of inquiry that may be placed somewhere or other on our ladder, at different stages of abstraction from the data of experience. If the results of their writing illuminate aesthetic experience, they are successful. Moreover, various routings are possible: the critic may only be interested in uttering true statements about works of art. If he does, he succeeds at his task. If any doubt is expressed about this success, he need appeal no further than to the careful self-criticism that is the task of metacriticism.

The criterion for success at metacriticism is the correction of criticism. And not all metacritics need appeal to theory or, indeed, to any fact other than a first-order experience of aesthetic objects and their understanding of a prior critical pronouncement.

The movement from metacriticism to theory is made by the metacritic who is interested not only in experience and criticism, but also in examining the consistency of his own principles of criticism with other possible bits of knowledge concerning the world. Or alternatively, some theorists, such as Santayana, noted above, may readily bypass criticism in the sense of evaluation and proceed directly to primary aesthetic experiences as data of their observations for the construction of a theory. And here too the inquirer may be satisfied with the hypothesis he has formed and confirmed by future, past-predicted observations. Either his inquiry will be satisfied with the confirmation of his hypothesis; or he may, in a moment of reflection, seek to check his theoretical inquiry on the basis of a general methodological evaluation of alternative hypotheses. Beyond this, there would seem to be no reason to go. The boundaries of aesthetic inquiry are therefore established between the aesthetic experiences of creating and appreciating on the one hand and the methodological clarifications of metatheory on the other.

V

We may now proceed to our third and last objective, viz., to indicate the relevance of aesthetic inquiry as heretofore defined for future educational research. There may be a great gap between the abstract possibilities of research related to the production and consumption of aesthetic objects and the actual building of a curriculum that would embody all, or indeed only some, of the positive results gained at any stage of inquiry. It does not follow, for example, that one should or should not do anything in the schools even if our own description of the scope of aesthetic inquiry be precisely what we claim it to be. Educational theorists must make some assumptions concerning the nature of their own activity in order to be able to use the results of any inquiry. In other words, nothing follows from our descriptions of the possibilities of aesthetic inquiry (read "research") unless the educational planner makes the prior decision to include aesthetic materials in a particular school setting. Such a decision must be made in view of the restrictions on the amount of time available for formal education in anything at all and the relative weights one places on the various subject matters to be chosen for actual formal instruction. In our time the relevance of mathematics and the physical and social sciences has been clearly established, and only recently has much attention been given to weighing the alternative of artistic and humanistic concerns.[38] But to what extent can it be truly said that our curriculum includes courses of related instruction in the creation and appreciation of art, in art criticism, in evaluation of art criticism, in aesthetic theory and methodological interpretation of such theory? And what factors, outside of an ignoration of the scope of aesthetic inquiry, militate against instituting such programs?

---

[38] The Central Midwestern Regional Educational Laboratory, Inc. sponsors as one of its continuing projects curriculum research and development in aesthetic education. The first phase of this project, directed by the late Professon Manuel Barkan of Ohio State University, has produced *Guidelines to Curriculum Development for Aesthetic Education*, copyright 1970 by CEMREL, Inc., St. Ann, Mo., Wade M. Robinson, Director.

It could be argued, of course, that one level of aesthetic inquiry—the theoretical, in which having aesthetic experiences is considered in the light of their effect upon the overall quality of human life—may be relevant to the educator's decision to include or to exclude subject matters which would constitute the basis for aesthetic education. But once again it is obvious that his choice must be made on the basis of a comparison between the relative capacities of various disciplines to achieve the overall aims of society in promoting education; in this sense, the educator must determine the maximal conditions for improving the quality of an individual's life or the fabric of the society at large, and his choice of subject disciplines will be directed thereby. Our explanation of the range of relevant aesthetic inquiry has therefore only what may be called a "contingent necessity"; that is, contingent upon the educator's decision to include aesthetic materials, it becomes necessary for him to consider both the scope and the relevance of possible aesthetic inquiries, each of which must itself be relevant to the facts of aesthetic experience.

It is our hope that this contingency will be met and that more educators will be prevailed upon to make the necessary assumptions. If they do, our explanatory "ladder" indicating the structure of aesthetic knowledge may be useful; indeed, if we are right, it must be used to lay out the actual subject matter for formal instruction, to wit, training in the making of works of art, in art appreciation and criticism, in metacriticism, theory, and metatheory. But such inclusion need not be presented at all levels of instruction. At the present time, for example, courses in creative activity are found useful to reinforce the aims of courses in art appreciation. These may be offered at any level and must be pursued with methods adequate to the subject matter being taught and to the ages of the students. And it goes without saying that only collective research will establish the measure by which individual classroom methods are to be judged.

An example of an educational theoretical inquiry into actual pedagogical methods is to be found in *Improving the Teaching of*

Art Appreciation, a report on the collective efforts of a research and development team whose purpose was to produce a model for the improvement of instruction. This research was funded by the U.S. Office of Education and carried out at the Ohio State University.[39] Let it be noted here, however, that these researchers were doing inquiry into the pedagogical dimension, partly descriptive and partly axiological, of aesthetic education. In their practice the whole range of relevant aesthetic inquiry constituted a subject matter and was judged in terms of its applicability within an actual classroom situation. The areas of expertise involved were represented by a historian of art education, a historian of art, a demonstration teacher, an art educator, a museum specialist, a philosopher of education, and a philosopher of art education (with critical responses by a philosophical aesthetician, an academic critic, and two practicing public school teachers, one of whom was also a city supervisor of art instruction). Each expert brought to bear what in his judgment was the best knowledge of both the subject matter and methods currently being employed in art education. A similar inquiry on a larger, more ambitious scale was undertaken by the Central Midwestern Regional Educational Laboratory in its aesthetic education project and is currently in phase two of its development.

In reviewing our own research, an educator may find it obvious that metacriticism and metatheory are of special interest at the graduate level of aesthetic education, and primarily useful as tools for the training of teachers of teachers in the so-called triple-T programs. More instruction must be provided in metacriticism in colleges of education if better courses in criticism are to be provided in the public schools; more courses in methodology and metatheory must likewise be provided there if, on the lower levels, we are to be able to supply defensible methods and theories for all those students whose experience within the educational process leads them to be skeptical or overly dogmatic about what they are learning.

---

39. See *Improving the Teaching of Art Appreciation*, Report on Cooperative Research Project Number V-006, U.S. Department of Health, Education, and Welfare, Office of Education, Bureau of Research, David W. Ecker, Project Director (Columbus, Ohio: Ohio State University Research Foundation, 1966).

There remains a last relevant consideration which has heretofore been left implicit in our analysis. The two formally methodological inquiries—the structure of aesthetic knowledge and the theory of aesthetic educational practice—supply nothing more than the formal limits of relevant research in aesthetic education. The present research requires constant reinforcement by the work of individual investigators who supply the material conditions—or, if one prefers, the facts—that give substance to the foregoing analysis. In this way, it was not our intention to replace the research of the workers in the various fields, especially not that of the artists themselves, but only to supply an interpretive schema for putting the fruits of their most successful endeavors into schematic organization, the better to be applied in schools of education and ultimately in the classroom behavior of teachers and students.

CHAPTER XII

Human Nature and the Scope of Education

DAVID HAWKINS

I shall be concerned in this paper with three postulates or conditioning hypotheses which have underlain the classical theory of human nature. My particular focus is the central place of human educability in the theory of human nature. In each case the postulates in question are relevant to education and in each case their significance has been blurred. For there is a wide array of plausible belief which appears to contradict them, and in any event they have never been consistently applied to the estate of childhood. On the other hand, there are serious difficulties and misfortunes connected with their overt denial. I propose to examine some of the fruitful consequences of these postulates, properly formulated, for a theory of human nature which gives place (as the classics do not) to education and development.

What is true of scientific theory generally is conspicuously true of the subject matter we deal with here. In the short run only a part of scientific theory is determined by empirical evidence supporting specific, testable generalizations; a part consists of some framework of categories and beliefs *brought to* the gathering of evidence and linked by boundary assumptions which suggest the kinds of generalizations to be sought and tested and the sorts of inferences these generalizations will sustain. In the long run it is not possible to justify a sharp and unchanging boundary between such relatively a posteriori and a priori components of knowledge. Research not guided by general theoretical formulations and arguments is likely to be inconsequential. Theory which is not linked to empirical research, to some form of practical experience and discrimination, is empty and cannot progress. The second of these propositions is widely recognized but the first is less so. In the "developed" sciences, like physics, theory is acknowledged to have sway. But in "undeveloped" fields like ours, it is often believed that theory is a sort of idle luxury or even a fraud, purporting to lay down encompassing generalizations which in fact have not been "empirically established."

On the contrary, it is in just such fields as ours that theory is of the greatest importance, though much neglected. For we do in fact already know a great deal about education and development; the problem is rather to marshall that knowledge in a coherent usable form. Empirical investigations conceived within such a framework of theory can react upon, can modify or extend knowledge. They cannot amount to much without it.

These remarks are by now only slightly controversial in the philosophy of science and may be illustrated by numerous passages from the history of astronomy or physics or biology. But let me come back to the particular categories and postulates I wish to discuss for the theory of education, picking up historical analogies as they seem helpful. For the sake of the orientation and controversy they suggest, let me use traditional labels. The postulated character of human nature is implied by three sweeping propositions: men are equal, men are free, and men are rational. These propositions have been often reasserted and argued over. They are the logical backbone of modern political theory as formulated in Hobbes's *Leviathan*, Rousseau's *Social Contract*, and Hegel's *Philosophy of Right*. The major consequences, common to this tradition, are the following: Without equality among men there is no profound *problem* of politics; the obviously superior will rule. Without freedom there is no deep *problem* of how the rulers can rule; they will rule by force or by psychological manipulation. Without rationality there is no alternative to force or manipulation, no meaningful debate over ends, no *problem* of consent or consensus. I mention these consequences because my own use of the postulates parallels them. Without them I cannot even define what seem to me to be quite central problems of educational theory and research.

In this specific context, however, my general remarks *are* somewhat controversial. It is usually believed that the classical theory (which, with emandations, I follow) is not descriptive but normative—that is, *morally* normative. For men are, often and profoundly, unequal, unfree, and irrational. The classical theory is taken, therefore, as only a hoped-for view of a case which, it has often been held, has been shattered by new discoveries about man and his career—discoveries in psychology, anthropology, literature, history; and so by a new and more sober enlightenment which penetrates through the optimistic superficiality of the eighteenth-century Enlightenment.

But the contrast between the normative and the descriptive is changed if we take account of the requirements of *descriptive theory*. For theorizing has its own norms—of self-consistency, of completeness, of definition and organization, of adequacy to subject matter, of fruitfulness. Because of these requirements a *good* theory outruns the "hard" evidence available to support it and if it does not, it is not a good theory.[1] But the basic postulates of a theory are not therefore to be taken as *mere* assumptions. They are supported by the kind and degree of organization they induce, by the new phenomena they lead to or highlight. They may be put forward wishfully—if so, the wishes that lead to a good theory are well rewarded!

I believe that a careful reading of the classical writers will show that their "normativeness" was primarily of this kind. The founding spirit, Hobbes, was first and foremost a theoretical physicist who sought to give the study of human affairs the same kind of logical organization which he and his scientific contemporaries saw developing from the work of Galileo and which reached its first great culmination in the work of Hobbes's much younger contemporary, Isaac Newton. I believe also that extensions of the classical viewpoint will prove their vitality for today and are sorely needed. So I wish to be judged on my position by the criteria of theory. But I do not wish thereby to conceal any of my own value judgments along the way. If the theoretical framework being developed is any good, it supports these judgments—and if it is not, then not.

The Educational Cycle and Its Closure

I begin by noticing a formal constraint upon educational theory which derives from the purposes it serves and which therefore determines its range or scope. A theoretical structure which fails to meet this demand is not thereby incorrect; it is simply inadequate to meet certain sorts of demands either correctly *or* incorrectly. In the classical writings this constraint is satisfied, but with equivocation and perhaps some embarrassment: *Adults* are equal, free, and rational; infants and children are, so to say, noticed in passing, but that is about all.

---

[1] Richard B. Braithwaite, *Scientific Explanation* (Cambridge: University Press, 1953), chaps. 3 and 4.

At its simplest, the demand is for a recognition of the full cycle of cultural self-reproduction. Thus persons must be describable both as learners and as teachers, as both subjects and agents of education. Childhood is no more transient than adulthood, they are both here to stay. In the cycle, childhood is an end as well as a means or, in other language, a sufficient condition of adulthood as well as a necessary one. In order to have a theory which is relevant over the whole cycle, its basic postulates must be formulated so as to obtain over the range from infancy to adulthood. In their ordinary usage the concepts of equality, freedom, and rationality are explained by reference to supposed characteristics of *adult* behavior. And among many of us, as among some of the classical writers, these characteristics are denied infants or children, or affirmed of them only in minute degree. But human development is not simply quantitative, and measures of it (such as the mental age of intelligence tests) are averages over many sampled dimensions of behavioral development. I do not claim to have found that formulation of the postulates which is invariant with respect to human age. Here, I am only concerned to state that this sort of formulation is a logical necessity for an adequate theory based upon such postulates. Without such a formulation childhood is left in a theoretical limbo, as it was, for example, by Locke or Hobbes. With such a formulation in hand we can then begin to trace those qualitative features of human development which will show completely what the postulates imply for each developmental phase. I shall later examine one attempt at invariant formulation, that of Jean Piaget. One obvious implication of an invariant formulation might be that children are sometimes teachers, curriculum specialists, and instructional consultants—with adults as their pupils.

The formal condition of the cycle is an obvious one to impose, but it is not trivial or always obvious in its logical consequences.

Economic theory, in a sense, began with François Quesnay and Adam Smith because they imposed the condition of circular flow and were loyal to its consequences. Once seen, it was an "obvious" condition and (at least in a strong first approximation) factually indubitable. But it led to unfamiliar conclusions about the health and wealth of nations. It widened the geographical and temporal scope of economics and through the circular linkage of production, distribution, and consumption gave rise to a kind of logical closure, so that inputs to consumption reappeared as outputs of production governed by a law of value—in effect, a conservation law analogous to those laws that were to appear later in nineteenth-century physics.[2]

The requirement that a general theory must apply over the whole cycle of the processes it studies is obvious in the case of biology. First, there is the cycle of interactions between an organism and its ecological niche. Second, there is the specific cycle of reproduction, necessary also to the study of evolution. As in economics, some implications of this requirement were realized only belatedly —in the time of Darwin; and that does not end the story. Faced with the problem of the origins of terrestrial life, biologists have only very recently realized that self-reproduction must have appeared before life itself, in the molecular milieu of the primitive earth. For no living thing without that capacity could have lasted long enough to acquire such a central and complex system of processes. Indeed, human biology implies in the same way that the basic processes of education must have evolved among our prehuman ancestors. For on a purely biological level our species is already apt for education and cannot survive without it. We are in a sense invented by education, not education by us.

The self-reproduction of a culture, like that of life or human economy, implies an essential continuity. As the caterpillar becomes the moth, children grow up and replace us who argue about schools, plan curricula, and teach. Such conditions are met in the real world, but some theories fail to meet them. Embryology has not always acknowledged the womb, or the womb within the foetus. So neobehavioristic theories with their S-R schemata have no language for the conditioners qua persons with methodologies, arguments, and truth-claims. Conditioners are *agents* (of a sort) for which the descriptive metaphysics of behaviorism has no place and which it can neither acknowledge nor deny, but can only exclude from consideration. In its methodological aspect behaviorism makes the cut between investigator and subject in such a way that the former is only present by implication. He talks about himself and his peers in a language of agency and choice which cannot be mapped into his own official theoretical language. This state of affairs violates the condition of cyclicity.

---

[2] I have worked out some of the logical details in *The Language of Nature* (San Francisco: W. H. Freeman & Co., 1964), pp. 333-41.

To accept the continuity of the cycle, one must impute to infants and children, with such transformations as development may explain, *their* methodologies, arguments, and truth-claims. Such a failure is no criticism of a theory within a more limited scope, though it does suggest criticism. Mercantilist economic theory had its range of validity also. But such a failure demonstrates theoretical inadequacy—in our case to the function of education. Nor does the recognition of failure lead automatically to a better theory, though it may again be suggestive. Solids and gases are pretty well understood today, but liquids are still a mess. So, for example, the large change in viscosity as gases condense to liquids is still a mystery —but the demand which makes it appear a mystery requires a theoretical framework which will itself weather the change of state.

The adequacy of a theory to such demands is, of course, a relative matter. The closure which education requires is different from that in biology. The continuous change from egg to human infant marks a conceptual discontinuity; this gap is of real but peripheral concern for education, the continuity of which is social rather than biological. Genuinely biological variables are important as boundary conditions for education but it is not defined in terms of them.

EQUALITY

The ground level of educational theory takes account of persons, of integral human beings as children and as adults. A postulate of equality in this context has a logical function similar to that which it has in our classical political theory. That is, it requires that certain kinds of de facto *inequalities* have to be explained as lying *within* the scope of the theory and not as biological boundary conditions. In discussing the meaning and validity of this concept I am not concerned to make moral judgments or to "argue for" equality (or freedom or rationality). In a scientific theory (as distinct from a hypothesis within a theory) a postulate is a proposition set forth as true. On the other hand, the import of a postulate, what it does to or for a theory, is of great importance in determining how to formulate it and how to exclude irrelevant connotations. Morality does imply certain factual egalitarian presuppositions, of course. There is a widespread weak-kneed concession that one should believe *in* equality and treat persons as equal, though of course one is not supposed to dream of believing there is any truth in it. Here, surely, where good theory is at stake, one wants to get the horse back in front of the cart.

The postulate of equality has the logical function of a symmetry condition. How such principles function in science is different in different cases but the logic is the same. To assert that entities of some class are identical or indistinguishable or equivalent in certain respects is to block certain sorts of explanation and to invite others. Elementary particles group in the same few identity-classes everywhere and they therefore do not per se explain the rich variety of nature, which is structure and history—not mere building blocks. The statistical equipartition of energy both blocks and redirects an enormous range of explanations in physics, starting with specific heats and Brownian motions.

Thomas Hobbes's political theory shows very clearly the logical import of the postulate of equality. His famous "State of Nature" is a thought-experiment based on this postulate. In the presence of limited resources for life and in the absence of government, men are inevitably competitive. Because they are so, they are dangerous to each other. In fearing this danger and wishing to forestall it, each man becomes more dangerous to his neighbor, and the situation must degenerate into "the war of each against all." Men are not so unequal in power or ability as to give anyone a natural dominance over others; and if one be for the moment more powerful, a coalition of others will have the ability to destroy him. Equality need not be absolute; the postulate need only be close enough to the truth to invalidate any argument which says that the natural superiority of some men is sufficient, in the "state of nature," to create a stable order based on dominance. In the absence of government, of civil society, men are not good or evil; they are simply determined by self-preservation to be enemies. Nor are they irrational; on the contrary, their very rationality, under these conditions, is what makes them mutually dangerous. If you were put down in such a world, its constraints would soon bring you to be an enemy of others, and like theirs your life would be "poor, nasty, brutish and short." Hobbes has often been criticized on the ground that such a state of nature has never existed, but this misses the whole style of his argument. Hobbes's theory is a mathematical model which in fact and even in some detail anticipates the von Neumann theory of games. The state of nature is a many-person, non-zero-sum game, but a game of life and death. Coalitions may form but they will be essentially unstable, for every power group which forms will be threatened by others which transiently form against it and by offers of new coalitions which divide it. As a result, all genuine possibilities of cooperation are stultified and men realize, on the average, a far less measure of security and livelihood than nature would allow them otherwise.

The "Social Contract," by contrast, is an alternative model. In this model all men, being equal and rational, realize the source of their misery and create by rational consent a sovereign person or group to whom they surrender their power as far as possible. This sovereign (preferably an absolute monarch) is not party to the contract but remains in the "state of nature." Self-interest, not an agreement, guides him to enforce civil tranquility, the rights of property, and the rules of contract. Under these conditions his subjects will be without fear of each other and in self-interest will cooperate with each other and will secure from nature their maximum welfare. Individuals may deviate and be punished, but large groups will not unite to violate the contract unless the sovereign fails to rule well and thus jeopardizes his own security.

There are many weaknesses in Hobbes's theory from the point of view of realism and adequacy, but it is a coherent theory and a sort of first (or, better, zero-th) approximation to the truth. I might mention that at the formal level of proof they sought, von Neumann and Morgenstern[3] were forced to stop far short of the general *n*-person game and its many alternative "solutions"; and needless to say, Hobbes's arguments are plausibility-arguments, not proofs. It is worth noticing that the later authors also adopt the egalitarian postulate; they do so precisely in order to show how social inequality can then arise *without* appeal to innate differences. Without the postulate their theory would be useless. And so it is with Hobbes. In the postulate of equality he blocked any theory of aristocracy based on the belief that we breed *natural* rulers; the postulate then leads to an alternative explanation of transient inequality in the chaos of the state of nature and also to the possibility of civil society having its investment of coercive power in a governing individual or body. The contract is only possible as an agreement of equals and only stable when governing a society of equals. (For the same reason the choice of rulers is theoretically arbitrary.) Others of the classical political philosophers used the postulate differently, but the logic—simplest and most daring in Hobbes—is always one which brings the origins of inequality *within* the sphere of human control or at least (as with Rousseau, Hegel, and Marx) within the realm of historical-political development.

---

[3] John von Neumann and Oskar Morgenstern, *The Theory of Games and Economic Behavior* (Princeton: Princeton University Press, 1944).

In education the postulate has sometimes been so formulated as simply to deny educationally significant congenital differences. I do not know any strict formulations postulating absolute equality, but many of us are disposed to believe, in the spirit of Hobbes, that individual or intergroup congenital differences with respect to education capacity are minor compared to those induced within society. From this point of view, unequal results of educational effort cannot be explained by differences of natural endowment, except perhaps for a necessarily small fraction of the population suffering gross constitutional defects. Such differences are therefore subject, actually or potentially, to educational control.

But rather clearly that version of the postulate is wrong. The cells of a single animal (disregarding occasional polymorphism) acquire their morphological and functional diversity as a result of embryological differentiation; my nerve, bone, and liver cells are identical genetically, but they differ by "education." But whereas cells of the same organism have the same heredity, congenital differences between organisms of a species are real and of great multiplicity. A large part of this intraspecific multiplicity (and conceivably, with the exceptions noted, all of it) may be of no great educational consequence. At some level of approximation this is certainly true. Hobbes spelled this out very crisply, as did Descartes.[4]

The purpose of scientific theory is not to *deny* differences but to describe them correctly. Human congenital differences, negligible at the level of a zero-th approximation theory like Hobbes's or Descartes's, have a positive importance for the concerns of education, and I therefore propose a different interpretation of the postulate of equality. If the purpose of the postulate is to place differences where they are relevant, then it seems clear one would do well to emphasize that persons of one very distinct species biologically differ constitutionally in many respects that are potentially important for education. On the other hand, there are no genes for algebra, for painting in oils, for operating linotype machines, or for Chinese calligraphy. Our glory and trouble is not specialized adaptation but generalized adaptability, and this is a universal characteristic of our species, common to almost all of our gene combinations. The niche of algebra and the niche of calligraphy are inhabited equally by many genetically different sorts of persons—if not, these niches would not exist. Nativistically conceived, human "abilities" are not Mendelian traits or simple functions of such traits. They are, in all likelihood, redundant complexes of such traits. Coupled thus to genetic variety is the species-characteristic of generalized adaptability, a certain kind of functional redundancy which characterized our capacities for learning. Put otherwise, the condition of cyclicity in the transmission and evolution of culture implies at most a very weak relationship between individual genetic variety on the one hand and cultural endowments on the other. This is part of the *meaning* of generalized adaptability.[5] To be a human being at all is to be biologically within the reach of many genera and species of human careers.

---

[4] Thomas Hobbes, *Leviathan*, chap. 13; Rene Descartes, *Discourse on Method*, Part I. From time to time the old beliefs of genetic determinism are brought out again. Niels Bohr was once told of two identical twins separated early, one becoming a farm worker and the other a university graduate. After thirty years they had the same IQ! Bohr's characteristic response was: "Obviously, they must have been very different to start with!" Of more serious concern is the periodic revival of claims concerning racial inequality, usually authored by brains from my own genetic stock. I suggest that *we* had better settle now for equality before there is an adverse change in the betting odds.

Studies and arguments such as those advanced by Arthur R. Jensen in the *Harvard Educational Review* (Winter 1968) share with the pure equalitarians the belief that "nature" and "nurture" are additive variables, so that one can say "intelligence is X% heredity and 100 − X% environment." Agricultural investigators have for a long time understood, even in their much simpler field of study, that this is only a limiting case and that *complementary* relations are more significant than additive ones. Cf. the analysis by Richard C. Lewontin, *Bulletin of Atomic Scientists* (April 1970).

Let me use and sharpen an analogy. Most of our machines are the very antitheses of our own adaptability. They are inflexibly specialized beyond anything in the biological domain, and even our great ingenuity can readapt them, for the most part, only at the slow rate of capital obsolescence. We may, therefore, seize upon the computer as the first machine having generalized adaptability within a domain—even though that domain is still a very artificial one. The theorem is a theorem of closure.[6] Any machine which satisfies certain simply stated formal conditions of operation can thereby do *anything* that *any* other such machine can do. Regardless of componentry and mode of operation, any computer can be programmed to operate as a replica of the operation of any other computer. If A has whatever specialized components and style of computation B lacks, then a suitable *description* of the former supplied to B in its program causes B to function in a way that "empathetically" duplicates A's operation. In this basic qualitative sense, all computers are equal. Of course the theorem says nothing about time or energy or cost, and it says nothing about what these universal computers as a class cannot do. Nevertheless, it establishes a sharp, if limited, definition of generalized adaptability, a domain of equivalence among machines across a very wide range of "constitutional" differences.

In our case the equivalence is of a different kind, of which

[5] This redundancy is implied also by the fact, underlined in the writings of J. B. S. Haldane, that man has the widest variety in his gene pool of any animal except the dog. See "Human Evolution," in *Genetics, Paleontology and Evolution*, ed. Glenn L. Jepsen, Ernst Mayr, and George Gaylord Simpson (Princeton: Princeton University Press, 1949).

[6] John von Neumann, "The General and Logical Theory of Automata," in *The World of Mathematics*, ed. James R. Newman (New York: Simon & Schuster, 1956), pp. 2070-98.

"computing" is only one rather special component, and we do not have the sort of theory to formulate a formal proof. The proof lies in the historical fact that we *are* the matrix of cultural evolution, that cultures are transmitted and evolved with no close dependence on our internal biological variety. Among other manifestations of our generalized adaptability we can, within wide limits, organize ourselves, each constitutionally in his own way, to perceive, to plan, to understand, and to enjoy, *each like each other.* Such an "empathetic" equivalence is a presupposition[7] of communication and of society.

But this is only the negative side. The old egalitarianism leads to one correct conclusion. Differences in educability are not part of the ground plan of our species. Differences are minor, though open, as we well know, to narcissistic elaboration. But there is a stronger objection to this viewpoint. One remembers the French joke about the difference between men and women. About individuality one wants also to say, "Vive la difference!" And this for theoretical reasons first, not moral or ideological or aesthetic reasons.

First, we well know that if any group of individuals are constrained to follow an identical pedagogical track, they will learn unequally, with a large variance in the particular learning rates which can be defined for just that track. The pure egalitarian assumption that children are indistinguishable at birth requires that this difference be explained as a difference in "readiness" associated with earlier experience. The single track remains unquestioned, even by most of those armed with slogans of "individualized instruction," which puts children to the book, or the machine, as though it were more adaptive than they.

But congenital variety among persons, modified by early experience, is not single-track; like biological variety in general, it is many-dimensional, its graph is a profile. And whereas a single well-defined curricular track will spread children out in a long line of march, there is also a variance in the learning abilities of a single child along alternative tracks, assuming for a moment that these are made commensurable by leading toward a common goal. Thus, by a proper assignment of tracks in a way which complements congenital variety, the variance of learning rates can be reduced and with no decrease of any individual rates.

---

[7] Cf. Descartes, *Discourse on Method, on Reason and Certainty.*

An interesting modification of the postulate of equality which idealizes this fact and is not inconsistent with experience (yet is unconfirmed enough to be challenging) is that for each child there exists *in principle* some unique optimal curricular track, of such a kind that in a group where learning is thus individually optimized the variance of attainment will be negligible (except again for gross congenital disabilities). I hope it is clear here that I am not speaking of tracks as prefabricated, as though they could be defined in a textbook on "alternative tracks." Mathematicians will recognize that I am speaking abstractly of education as a function (or as a functional) over *two* domains of variables which are essentially nonlinear in their relations to each other. We cannot speak of an optimal track in general, except with reference to a particular learner, or of an optimal learner except with reference to a particular track. Of course, there are some universally bad tracks.

Such a postulate is not all that I am after, but it is a way station and it sustains the egalitarian theorem that human "inequalities" are only by way of exception *derivable* from congenital differences. For, although not widely practiced, appropriate provision for qualitative variety in learning styles *is* within the domain of potential educational control.[8] Moreover, the postulate has already the right logical form: human beings qua human are never indistinguishable, never identical to each other, even in respects important for learning and education. They are, rather, *incommensurable* in their differences. For the new postulate as stated leads immediately to a further theorem, which is that by a suitable choice of learning tracks (nonoptimal for some) it is possible to rank n learners in all $n!$ possible permutations of "ability." It moreover suggests a stronger theorem (depending on some plausible side conditions) that for each such permutation there exists some *single* track that would realize that very permutation of the $n!$ individuals. It would take $n!$ no-option schools to do it, in theory: to assign an arbitrary rank order to n children and then pick or design a school which would realize that order, as determined by some antecedently prepared final examination.[9]

---

[8] Many educational experiments sample children with statistical care, but the range of possible educational environments is sampled very poorly.

Thus, incommensurability excludes marketplace or IQ-dominated notions of unalterable inequality implicit in "the superior student," "ability grouping," and so forth. For these imply commensurability, the belief that "the more able" excel "the less able" along all possible tracks. In this connection it should be emphasized that incommensurability implies that individuals *can* be compared and ranked in many sorts of ways. It means that such comparisons are vector rather than scalar in type. It implies that in general one individual does not excel another in all relevant dimensions, does not, in mathematical language, dominate him. Acceptance of the postulate would have the salutary initial effect of raising an eyebrow at *any* course of study or program directed to "the superior," "the gifted," and the like. There is no doubt that along the small variety of curricular tracks available in our schools conspicuous dominance is unavoidable; indeed, it is sought after in the uniform sequencing of curricular items and lesson plans—self-confirmingly so in the case of many sorts of "ability grouping." [10]

Since the postulate of pure equality or educational identity allows an explanation of variance in learning only as a function of previous learning, it is therefore likely to lead to an emphasis on "remedial" education conceived in terms of slowing down and catching up, e.g., of giving the "retarded" a "head start." Much of the recent effort to redeem the school education of the disadvantaged has taken for granted a particular school language track along which many children are disadvantaged, and concluded from this the need for a remedial priority upon spoken and written English, an emphasis which in many cases guarantees the very alienation which it is intended to remedy. But however much one would welcome in schools (not least for children of the ethnic majority) the addition and enjoyment of black dialect, or Spanish, or the Indian languages—all previously unused and characteristically forbidden—the translation of existing contents into differently colored capsules is a limited innovation, to say the most. The Andersen fairy tale could easily tell in three languages how the king's new clothes were described.

---

[9] Again mathematicians will recognize that I am illustrating a theorem about vector inequalities. Perhaps it couldn't be done. And perhaps it couldn't be done because we're not clever enough. General postulates are not easy to prove or disprove, and that is what makes the choice of them so critical.

[10] Although even after obvious gross inequities in early life, groups of children defined by ethnic and economic background still have extremely *uneven* profiles on conventional ability tests. See G. Fifer, "Social Class and Cultural Group Differences," *Proceedings of the 1964 Conference on Testing Problems* (Princeton: Educational Testing Service, 1965).

The postulate of incommensurability, on the other hand, takes children as congenitally varied rather than "unequal," and raises questions about the *differential* effect of earlier environment in relation to the kinds of learning it has supported or inhibited. It underlines the importance of local and dependent curricular and instructional choices, to make the curricular spiral tangent at many points to the individual lives of children, to the educative resources of *their* total environment which *they* know or can be helped to discover. (It underlines also the importance of teachers as professional persons who diagnose and plan, not as conveyors of preorganized material to the young and who only when worn need "retooling.") This proposition is no less important for the education of "advantaged" children; it is only at present less in the political focus.

A postulate of sheer inequality, of the natural dominance in abilities of some children over others (racially correlated or not), leads to a quite different emphasis, namely, to the existence of tracks or streams running more or less parallel in the early years and then diverging to appropriate terminals from which each child is to be emitted into his appropriate station in life. Such sortings are strongly self-confirming, a fact visible even to those who embrace dominance as a "hard fact" of life. Such persons are directed by their assumptions to a search for better sorting mechanisms, not to a questioning of their own presuppositions. Some political egalitarians are suspicious, on the other hand, of all proposals for greater internal diversity, of "pious talk" about the educational importance of choice and self-direction of children. The stereotyped image of the curriculum as a sort of "progressive" self-help cafeteria administered in the spirit of laissez-faire in which each child is somehow busy "fulfilling his own potentialities" may indeed properly be condemned as the benign work of an educational system geared to the preservation of social inequality—the malign one being a system of meritocratic groupings or streamings. A rigidly single-track system at least—so it is argued—guarantees "equality of opportunity," however unequally and ineffectively the opportunity in question impinges on the lives of diverse human beings.

But the meaning of incommensurability is that diverse children *can* attain to a common culture—a common world of meanings and skills, of intellectual tools, moral commitments, and aesthetic involvements. Individual development *can* complement individual differences, but only through a matching diversity of learning styles and strategies. Children can learn equally, in general, only as they learn differently. The more constraints there are toward single-track preprogrammed instruction, the more predictably will the many dimensions of individual variety—congenitally and individually evolved—express themselves as a large rank-order variance in learning.

But I described this statement of the postulate as a half-step only. For there is no necessary implication that the goals of education be concretely the same for all groups or all individuals. The idea of multiple tracks toward a common proximate goal is one way of expressing our postulate, but not one which exhausts its meaning. Experience and learning are many-faceted. Though two tracks lead to algebra (or calligraphy) for two children, the fact that one is optimal for one child and not for another implies that neither track is relevant *only* to calligraphy or algebra, that neither is relevant *merely* to one goal. A track is efficient toward one goal because it is linked to a child's existing interests, knowledge, and skills; being so linked it relates to and supports other goals as well. Algebra is valued *for* the curriculum only because it can be richly cross-referenced to many other domains of experience of form and analysis. Indeed, one can hardly defend its central place in the curriculum without such reference, or therefore defend the teaching of it, as is mainly the case, along a path which does not gain support from and give support to quite other kinds of interests and activities. No track toward algebra can be efficient which is a track to nothing other than algebra and from nothing other than text and teacher. The graph of the domain is not of a collection of mutually isolated tracks but a network of richly interconnected nodal regions, each of which *can* be strategically isolated at some times and for some purposes. But the mind can only isolate or block off connections when it can grasp them. Otherwise, isolation is busywork and we have that in abundance.

For those who would seek to predesign a curriculum by little programmed steps, implications of the network are onerous indeed. If a branching "program" allowed one binary choice per month and every monthly "package" or "unit" were designed to fit all previous choices, about a thousand tracks would be required by the end of the year, and by the end of twelve years, more than there are grains of sand on the planet. Fortunately, the world beyond the symbol is self-programmed, and so, in a sense I shall try to define, are learners when their capacity is not deadened but given active support.

Diversity of pathway in learning implies the network rather than the little racetrack (*curriculum*) and underlines the necessity in learning of constant choice and invention. Human beings are valued within a community for their useful differences moreover—as sources or resources of skill, of aesthetic expression, or moral or intellectual authority. It is not difference as such which we value, but individuality—the unique personal style and synthesis which interests us in each other as subjects of scrutiny, of testing, of emulation, or repudiation. Recognition of individuality completes what I mean by the postulate of incommensurability. The character which members of our own species possess—what we term individuality—implies neither dominance nor identity, but equivalence within a domain of relations sustained by individual diversity. If the old word *equality* should be used in this sense, it is the equality of craftsmen working at different tasks and with different skills, but with plans and tools congruent enough to provide endless analogies and endless diversions. Or, it is the equality of authors who read other authors' books but must each, in the end, write his own.

Educational traditions in their practice or their rhetoric may support or oppose such notions of individuality. I urge them as simply correct notions, receiving support by the same logic that combines diversity of endowments and histories with a complementary redundancy in these endowments—generalized human adaptability. In this way one allows that a common cultural flux is transmitted diversely, yet provides a currency of skill and knowledge accessible to all. Such incommensurability is no weak postulate. In allowing only relational or vectorial differences—differences of "profile" across many entries—this logic denies dominance, denies any single axis of inferiority-superiority. And if we are to urge or claim that we provide "equal opportunity," the postulate requires a providing of great diversity in that opportunity. By this test we have not gotten very far. Along that one narrow track which we do mostly provide, dominance must prove the order of the day.

Discussion of the first postulate has partly anticipated that of the second and third and must partly follow. The implication is clear, I hope, that in discussing match and mismatch between native endowment and learning-career we are talking about learning which is of educationally significant dimensions, irreversible, and, in any detail, nonreplicable, which has a time-scale that spans years or decades. Learning theories trimmed to the laboratory scale *may* have some relevance, just as the clinic may have. But extrapolations from the ambient of laboratory or clinic to education are almost unbounded, as is the shift in subject matter. To assess relevance here we must be independently acquainted *in* the education domain and not just extrapolate to it. To judge well of such matters we must be widely acquainted across the domain of children *and* of the variety of possible teaching situations. Professionalism in educational research ought to interpenetrate, far more deeply than it now does, the professionalism of teaching.

FREEDOM

The discussion of human freedom—"freedom of the will"—is familiar to most of us in the form of a debate that goes on, starting nowhere and ending nowhere. The focus of the debate is whether "it" does or does not exist, is or is not a meaningless concept, is required by religious faith or excluded by science, and so on. People seem in this connection to know what the word means, although this is usually rather doubtful. What I have in mind to discuss here is much more limited, fortunately. I wish to explicate how a postulate of freedom functions or can function *in a theory* of human nature in which the meaning of the postulate can be explicated by the logical uses made of it. As soon as the term *free-*

dom is looked at in this way, it loses some of its blank metaphysical debatability and becomes, instead, rather more consequential and interesting. In particular, the word *freedom* is not at stake. Or, rather, what is at stake is not *the* meaning of the word, but a particular meaning which is tied to some very central propositions about human nature—tied by these propositions both to learning and knowing on the one hand and to choosing on the other.

In the classical theories the postulate of freedom, as I call it, functioned in the same way as that of equality—as a preliminary blocking move which limited certain kinds of explanation of human behavior in order to make way for others. In blocking some explanations and allowing others the postulate lays the ground plan for a special *kind* of causality; it does not deny causality in general (as though "free" meant "uncaused"). It asserts the efficacy in human affairs of action with knowledge, of action guided by knowledge, including the special kind of action involved in making decisions and in seeking more knowledge. And on the negative side the postulate denies that explanations in such terms are universally reducible to other kinds of explanation, as, for example, those common in physics or biology.

As in the case of equality, so here, I am not mainly concerned with historical theories but only refer to them as touchstones for a body of theory which we are all at least vaguely aware of from our political traditions. There are interesting differences among the classical theorists which I shall not allude to. And I am interpreting them freely, though, I think, not at all wrongly. In particular, the problem of the reducibility of such concepts as "action with knowledge" to purely a physicist's or a biologist's language is a long story, and I shall cut it short by appealing again to the requirement of adequacy over the whole cycle for which educational theory must (like political theory) be responsible. All of us who debate matters of educational theory and practice are trying to act with knowledge, and we describe children as acting that way as well; it is our aim for them. Since our children will one day be carrying on our same effort in education (on a higher plane we hope), we have no choice but to accept this irreducibility of intentional, personal language, not as a stark metaphysical necessity perhaps, but as a logical requirement of adequacy in our theorizing.

Men can be overwhelmed by superior physical force; they can be damaged or destroyed by disease; they can be constrained by physical limitations or by other men. But within such limitations there is a wide range of human behavior—of a kind which is emphatically human—involving perception, memory, knowledge, and choice. The existence of illusion, misremembering, ignorance and error in reasoning is part of such behavior, of course, and facts of this order may in one sense be explained in terms of physical or biological boundary conditions. But such categories of fact exist only by way of exception. An error is only equivocally the effect of physical causes; the explanation of an error requires that it be recognized as an error first, before an explanation is offered. Since a correct explanation will require that the error be correctly recognized, *that* recognition, as well as the explanation, is an act of knowing. Thus all such privative concepts as error, illusion, etc., derive from corresponding affirmative ones and physical causes involved are not enough to explain such human behavior.

In the explanation of behavior in the normal range there is, indeed, a very essential reference to the phenomena of men's physical and social environment, but these phenomena are referred to *as* taken account of—perceived, recognized, enjoyed, compared, predicted, generalized about—and in that case these acts and their consequences in further action cannot be explained merely *by* the phenomena in question; the explanation is also in terms of how they are or are not taken account of and made use of.

Hobbes and Spinoza offer a general framework for this kind of explanation in the way the postulate is stated, connected with the idea of self-perpetuation. It is a "law of nature" analogous to Galileo's law of inertia that human beings will act to preserve and stabilize themselves. This is not something they *choose* to do, but is implicit in *every* act of choice. Learning and knowing depend not upon external impressions but upon an "endeavour outward" (Hobbes) or a "conatus" (Spinoza) which simply expresses the individual's nature, thus his self-determination or freedom.[11] Insofar as the individual organizes his experience in accordance with what he already knows, he extends his knowledge and his "power of acting" (Spinoza).

---

[11] Hobbes defines freedom in terms of the absence of external constraints upon this natural functioning of human beings. Spinoza criticizes Descartes for allowing a kind of irrational "freedom of will." The significant freedom in Spinoza's system is action which follows from adequate ideas. The power of the mind, he says, *is* knowledge. See his *Ethics*, Book II, prop. XLIX; but one must study the whole work. Sometimes this view of freedom, which

For both Hobbes and Spinoza emotion has a central function in this process; it monitors the gains and losses of this internal organization and power—a sort of feedback loop, as we would call it, which steadies the process. Thus, joy is not *what* men seek; it is the *evidence* of success in what they seek, as sorrow is the evidence of failure. "Affective education"—in a current mind-numbing jargon—is simply good education, education for self-direction and self-government.

Thus, freedom is something that men have by degrees; there is an innate core of freedom—a capacity for using knowledge in deliberate choice—which expands with experience and knowledge, so that it is not possible to draw a sharp distinction between self-preservation and self-development. In preserving their being men learn and, in learning, both the thing preserved and the means of preserving it are enhanced.

A contemporary expression of a very similar kind is to be found in the general theoretical framework of Piaget. This framework has, of course, a good deal more substantive psychological content built in and around it than one can find in the earlier thinkers (though I wouldn't be sure it is more penetrating than Spinoza's). But neither of the earlier writers knew how to be explicitly and

---

Hegel also made his own, is stated paradoxically. By the measure of the adequacy of one's knowledge and insight, one's actions become free of accidental local causes. These do not account *for* one's action, they are taken *into* account. In a compressed formula one expresses this by saying "freedom is self-determination." But the formula is too condensed. Freedom is a special *kind* of self-determination, mediated by learning and choice. This is not an attempt to "escape" from the idea that human actions are explainable, but only to explain an important class of actions correctly. I have developed my own view more carefully in *The Language of Nature* (see note 2).

For those who wish to avoid the classics, a contemporary reexamination of these basic questions which avoids all those taboos called "anthropomorphism" is to be found in some of the literature of General Systems, particularly in the analysis of the concept of an "open" system which possesses a certain kind of stability in its interactions with a changing environment. See W. R. Ashby, "Principles of Self-Organizing Systems," in Heinz von Foerster, *Principles of Self Organization* (New York: Pergamon Press, 1962). The only "formula" which will predict the behavior of such a system is to process the same information it processes, in the same way. "I said, if I were a horse, where would I go? And I went, and there he was."

empirically concerned with the full cycle of life; as is true of most thinkers of the past, they were fixated upon the presuppositions of adulthood. Piaget in this vital respect has broken fresh ground.

So in my reinterpretation of the postulate of freedom I shall make use of the logical pattern of the Piagetian terms *assimilation*, *accommodation*, and *equilibration*, using it as an example of the way in which a rather well-developed scientific theory can postulate and make essential use of an explanatory scheme not reducible to patterns of purely physical or biological causality. In the next section I shall indicate what I think is an important source of difficulty or incompleteness in one part of this theory as it has been put forward, but this is not to put it aside.

An important, implicit part of the idea of assimilation (suggested by the biological analogy) is that of a preexisting match between sensory input and the system of active processes which receives that input. Ultraviolet and yellow light will equally enter the eye and be scattered or absorbed as physical energy. But the sensitivity of the organism to these two "stimuli" is radically different. A few photons of the former may cause negligible changes in a cell or two, while a few of the latter may trigger the release of cellular energy throughout the body many orders of magnitude greater. From a purely energetic point of view the two photons are not much different; the difference lies in the optics and biochemistry of the receptor and the organization of the associated nervous system. From a neurophysiological point of view light frequencies near that of yellow light are "stimuli," others are not. And even within that range much else is excluded. In normal vision we do not see photons nor do we see retinal images made by them; what we are tuned to see are minute variations or modulations of the energy-input to our eyes, those, namely, which inform us of things around us: color, figure and ground, spatial relationship. Out of the virtual infinity of patterns which are in fact conveyed by the light, we *see* only those which are significant to us. The visual system involves a complex response-manifold within which specific physical inputs are "stimuli" and outside of which they are not. At far higher energy intensities there are of course more generalized biological and purely physical responses, at the limit of which the organism is destroyed. The decisive characteristic of neurophysiological response is the high-order amplification of selected (and energetically very minute) physical inputs. It is this fact which constitutes the stimulus, not as energy or momentum, but as thermodynamic information, as entropy change, and *this* constitution is entirely determined by the existence and nature of the response-manifold. But that is not all. The investigation of the stimulus-manifold, of the way we describe what a particular stimulus *is*, is entirely a matter of inference from the corresponding manifold of responses. We may know much about the environment of an organism without knowing what aspects of that environment will be responded to or, therefore, what constitutes "stimuli" within it. "Stimulus" is a neurophysiological concept at this level, not a concept of physics, and not yet, except by analogy, of psychology.

If we speak of learning in this neurophysiological sense, we may use any one of a variety of unconscious conditioning experiments as examples, including those of visceral, heartbeat, and breathing patterns. But clearly we cannot extend this scheme to other kinds of learning, such as learning the Pythagorean theorem or learning how to produce visceral conditioning. Learning in the educationally important sense is another jump in informational variety. It is in the range of these evolutionary capabilities that human culture, education, and individuality have their roots. The locus of education is, as its focus ought to be, in the concern for the way this accommodative evolution takes place.

It is in this context, I believe, that we can most fruitfully discuss another aspect of the postulate of freedom, the voluntaristic "negative" aspect which suggests not only that human behavior is not determined by *physical* boundary conditions, but that it is somehow unpredictably elective, erratic, and (to a suitably unprepared observer) "random." Indeed, the combination of high selectivity and high amplification gain in all response patterns of the brain, and especially in those characteristic of human learning, is calculated to defeat any pseudophysicist's explanation, and no physicist would find the effort interesting. We know enough about animal organisms to know we are dealing with systems precisely characterized by *this* kind of unpredictability (their survival depends upon it). We knew it all along, of course, but neurophysiology is needed to answer (like an ungrateful child) the mechanistic traditions that cradled it in its earlier history.

But it will not do to accept the strictly physical unpredictability of behavior as a *necessary* condition of freedom, without at the same time examining what kind of causality we put in its place as a viable alternative. Even in the case of simple conditioning we must learn how to *define* a stimulus before we can link it to a response. The relevant dimensions of informational inputs are not those of physics but a mapping of them into a manifold of responses. *Physically* different stimuli will evoke the same response on different occasions, and *physically* identical stimuli will evoke different responses. Behavioristic psychology recognizes this in a variety of ways (response set, intervening variables, and so forth). Thus, if we try to predict responses to stimuli we must first be able to *describe* the stimuli in relevant categories.

The jump from purely psychophysical learning phenomena to educationally significant learning is defined, in part, by the condition that in the latter case the definition of a "stimulus" is an act of the learner and not merely of the experimenter or teacher. The intentional meaning of *to learn how* or *to learn that* implies not only the de facto acting of a stimulus (as described by an investigator or teacher) but also a *taking* of the stimulus by the learner as an instance of a kind (correctly or not, for the purpose at hand) and a relating of it, so taken, to other categories. The verb *to learn* in such a context carries with it a truth-claim which is ineluctable, and this obvious fact escapes the net of theories of learning cast in associationistic or "S-R" language. One key to a positive theory of causality in the human domain is provided by this concept of recognition. The boundary line between assimilation of information which is unconscious (as in visceral conditioning, and so forth) and that which is conscious lies here. Recognition is the gateway.

Just as a neurophysiological stimulus is constituted as a stimulus by the response characteristics of a sense organ and of the whole organism, so at a conscious level a stimulus is constituted as such by some act of recognition. At its most primitive, recognition is the perception of difference and similarity, a preliminary mobilization of resources for meeting any new situation. The mapping is selective, so that differences and similarities perceived must be understood as standing out against a background of potential differences and similarities which go unrecognized. Recognition in turn directs recall, the search routine by which past experience is selectively brought to consciousness, to the locus of discrimination or choice. Recognition and recall may directly terminate a conscious phase of experience—in Dewey's language, may resolve a problematic situation. These processes may even become so habitual in standard situations that they lapse into the automatism of unconscious performance, into the mode of behavior which is in a gross quantitative sense our predominant way of being. In this sense we condition ourselves. Driving a car, riding a bicycle, typing, even solving equations can be described along a continuum. At one extreme is the fully conscious working out from an essentially problematic situation in which behavior is being defined and established. At the other extreme is the fluency of highly perfected response which has become literally unconscious—in which focused self-conscious discrimination has disappeared and in which our conscious capacities are occupied elsewhere if at all.

The postulate of freedom, of self-conscious agency, requires a nonphysical and nonbiological language and it cannot, therefore, be tested by purely physical or biological means of discrimination. But it does have negative implications of such a kind. We know that a pipewrench does not "think," nor does a typewriter, a telephone network, or an on-line computer console. If we take the nervous system or, better, the whole human organism as our subject matter, its operation as a physical system is in nowise excluded from study by the postulate of freedom. What is excluded is that this study should lead to explanations of human behavior which violate informational conditions—e.g., a pep pill which would "cause" me to write a good book, or an operant-conditioning regime that would "cause" me to understand some of the troubles of education.

If such an analysis of the postulate of freedom is at all correct, then it becomes possible to inject a certain theoretical discipline into recurrent controversies over the aims of education. Today there is a resurgence of criticism of the institutional habits of our schools on the ground that children are treated as things rather than as ends in themselves, that the school conceives the process of education as one of "transmitting" knowledge, "developing" skills, "training" children (or teachers!) by routines of instruction which in fact minimize opportunities for significant choice and self-direction. Children are not conceived as coagents in the process of education, but only as patients, recipients, and this general line of criticism sees such practices resulting in boredom, alienation, "failure." The question thus raised lead to the next section, but one observation is possible here. Many of the critics who hold these views appear to operate with a simplistic notion of freedom, such that the implied cure for the troubles of education is simply a switch from the active-teacher–passive-child model of school to an active-child–passive-teacher model.[12] It is the apparent moral presupposition of such a view that active planful intervention in children's affairs is per se a stultifying indignity. It treats "freedom" as an innate capacity or one which develops best when least hampered by adult direction. It thus cuts off the whole range of active human intercourse in which the transmission of culture can enhance children's freedom, their aptitude for significant choice. The view repudiated does not trust children's choices at all, whereas the critic's view sees no need to cultivate and help organize them.

RATIONALITY

The postulate of freedom makes it explicit that the central sort of causality in educational theory is that of agency, of the human decision-maker, teacher, and learner, and adult and child. But "agency" as a disposition is a singularly blank one except as filled out, in theory, by reference to those special characteristics of human existence which can be attributed to it. In the philosophical traditions from the time of Plato and Aristotle the operative meaning of human agency has been associated with what I have called the postulate of rationality.

In the classics *reason* has two aspects. On the one hand it is a functional disposition or capacity; on the other hand it is an operational domain into which individuals enter but which is by its own nature objective and public. Private thinking may be more or less rational, but it may also be intuitive, idiosyncratic, wishful—irrational. But the valid communication of private thinking, its communication in such a form that others can accept it *as* valid, requires that it be mapped or translated into a form which is invariant to individual idiosyncrasy. The generalized adaptability of which I earlier spoke allows the creation of a public domain of thought and action, a domain invariant to the replacement of one individual agent by another. The development of rational capacity is in Piaget's language a kind of decentration which takes place along with the development of schemes of operations. Once thought is functioning within this scheme, it is per se rational. I think this is why the classical writers tended to restrict themselves to adulthood, to the "age of reason"; it was an implicit recognition of the developmental nature of the scheme. They did not study, as Piaget has tried to do, the way in which prerational thought at the same time prefigures the rational. In the conception of rationality which we now must seek, there must be certain invariants, not only between adults, but also along the life course of one individual growing up in association with others.

---

[12] For a careful initial discussion of what is involved here, see Anne M. Bussis and Edward A. Chittenden, *Analysis of an Approach to Open Education* (Princeton: Educational Testing Service, 1970).

Piaget's conceptual trinity—assimilation, accommodation, equilibration—purports to do this, although of course only as filled out by empirical developmental evidence and only by means of subordinate concepts such as operation, coordination, decentration, and so forth. If there is a single central theme here which can be a candidate for our generalized concept of "rational," it is the tendency, manifest from early infancy, toward the organization of perception and behavior in terms of a hierarchy of invariant structures which is open at the same time to a reorganization of those structures as still higher levels of organization may require.

Seen from the point of view of the individual career, this unending process, with the stages and transitions it achieves, is "the life of reason." This does not mean that we are conscious of everything that happens in it. On the contrary, our ability to review and characterize what we have achieved or failed at is possible only from a point of view outside of that achievement or failure, and is part of a *further* integrative process not yet retrievable by us for review. The conscious agent is like a weaver who works at local pieces of the pattern without being able to see the pattern of the new weft except as the parts already finished may help him to define it. But without conscious agency and choice the whole process would cease.

Thus we come back to the concept of the human agent. Whether we think of the agent in traditional theological terms or as a center of coordinative activities in an open neurological system (or both), the "epistemic subject" is a necessity of thought in our framework, and as such does not have to be justified by outside interpretation. Our concern here is with function, not substance. We must see ourselves and others as agents, not because this mysterious essence is a tangible object of our awareness, for, if it were, it would be only one object among others and would not serve its function. We need the concept for a reason analogous to the need for coordinate systems in geometry. What I once organized and stored I may now retrieve for a new use provided I can now define the perspective, the viewpoint, from which I then worked but which was not *then* available to me for definition. In this sense there must be more retrieved than was stored. An example from Kant concerns the transition from a succession of perceptions to the perception of their succession. I hear each successive sound of a bell, but that fact does not imply that I perceive their succession. The first sound is fixed in my "circulating memory." The second does not merely "recall" the first, as though they might be confused. The recall brings a tag with it, so to say, and the second sound is modified as being second, as thus *related* to the first. Kant called this a synthesis in intuition, and it is in a sense a prototype of all higher synthesis, perpetual and conceptual. Along with the perception of succession in such individual perceptions there is also, of necessity, a recognition of the identity of the agent from one time to the next, who can retrieve now what *he* stored earlier.

When we talk about a child's ability to make use of resources which he has earlier assimilated and transformed, we are being concerned with precisely that sort of synthesis which Kant made the basis of his famous *Critiques*. But we are not taking it as something automatically given along with the definition of "the human mind"; we are taking it as a capacity still incomplete and evolving. Thus *we* seek a causative role in the process, and if we do not, we are not talking as teachers or educators. It is from this point of view that the framework of Piagetian theory, though developmental, is incomplete. Piaget does not profess to be constructing a theory of education; he leaves uncontrolled all those variables which educators must seek to define and find ways of holding within their hands. But he does, along with a neglected philosophical tradition, help us to define what some of those essential variables might be and, above all, what they are not.

In pursuing this line of discussion, and in trying to suggest what these essential variables might be, I shall first review some leading ideas of Kant, and then of John Dewey's maturest writing. From this I shall try to suggest how we may begin to formulate some of the dimensions involved. This leads me to a critique of the developmentalist way of thought, which with its longitudinal perspective appears to ignore these modes of educational influence. The construction which begins to emerge from this critique is one which on the one hand implies a certain sort of diagnostic function by which a teacher tries to map a child's current patterns of involvement and levels of organization. It implies on the other hand the teacher's responsiveness, his aim to facilitate those sorts of assimilation, transformation, and synthesis which diagnosis indicates will match the child's current involvements and readiness—to make means available to the child which the latter does not yet quite know he needs. This of course has often been said. What is also familiar, in fact, though less often said, is that the core of this diagnosis is that "empathetic" adaptability of which I spoke earlier. A teacher builds a model of the child's performance within his own familiar ways of operating and thinking and stretches himself in the process. It is thus that he defines the child's strengths and weaknesses, the continuities and gaps. This means that his diagnostic description, though fully behavioral, is not "behavioristic." It is cast in the same language of agency and choice by which he communicates about his own thought and action.[13]

---

[13] Here we can join issue, from the point of view of educational theory, with Martin D. S. Braine, "Piaget on Reasoning," in a recently republished monograph in *Cognitive Development in Children* (Chicago: University of Chicago Press, 1970), pp. 35-54. Braine criticizes Piaget's "epistemological confusion" in supposing that he is investigating "logical processes" going on in the child. As Braine says, if you think something you are independently familiar with is going on in someone else's mind, then your job is diagnostic, not purely descriptive. In the reformulation he offers of Piaget, this whole scheme of logical processes is taken out of the psychological domain and put back into that of subject matter, what we learn *about*. This "epistemological confusion" of imputing an operational model to the child's thinking is just

In the traditional array of "faculties," reason's appointed role was that of guiding and unifying the operation of the more special faculties—understanding, will, and feeling. Thus reason in this global sense was not restricted to "intellectual" matters. Its scope is indicated by the range of Kant's three *Critiques*: that of *Pure Reason*, in relation to knowledge of the natural world; that of *Practical Reason,* in relation to planning, decision-making, and morality; and that of *Judgment*, which could also be called the critique of man's aesthetic capacities. In the first *Critique*, the understanding is effective in relation to the concrete particulars of the phenomenal world as these are perceived and organized within the framework of the categories of knowledge. Reason by contrast stands above and guides the understanding; it is reflected in our concern with the *career* of knowledge, with the *ideals* of organization of knowledge, and with the *norms* which guide its pursuit.

In the second *Critique*, reason, by virtue of the role it plays, has a more commanding position. Moral choice, as Kant describes it, is a sort of adversary procedure out of which conduct is forged. This procedure, which forges out new patterns of conduct, takes place "before the bar of reason" with its categorical demands for consistency and finality in choice. But these demands are not coercive; reason is effective only through the respect it can command, and only in the sense in which respect can be commanded: through the enjoyment of that autonomy which reason alone can bestow upon the action of the will, which otherwise remains the captive of unexamined habit or impulse.

The role of reason in Kant's third *Critique* is less explicit and is developed less formalistically than in the first two works. From some modern points of view, though not in terms of historical influence, it is the most important of Kant's systematic writings. In both of its two parallel essays, one on aesthetics and the other on natural teleology, Kant is examining what he calls the *reflective* judgment, under circumstances such that the validity of scientific

---

what a good teacher is constantly trying to do. A teacher is busy being a participant-observer of a child's learning and must therefore read behavior as indicative or diagnostic of what is "going on"; it is only through such a reading that he can summon up the resources which will match and assist the child's. Piaget's "confusion" is what makes him useful to us. Braine would make the same criticism, I think, of Kant.

inference or the finality of moral choice has no counterpart. In the domain of art Kant examines the faculties, as it were, *at play*, detached from theoretical or practical ends. This interplay of perception, imagination, and understanding, prized for its intrinsic affective content, enables Kant to study the operation of human analytic and synthetic capacities—so to say, the *style* of a rational being—without his usual concern to "justify." What is laid bare in the study of phenomena of aesthetic experience or of man's reflective effort to find aesthetic satisfaction in his relation to the world about him is characteristic of all experience; but elsewhere attention is distracted from this subjective aspect by preoccupation with objective content.

In what is in many ways his most important writing, *Art as Experience*, John Dewey covers or re-covers this very same ground, but from a different starting point and in a different style. Although he mentions the *Critique of Judgment* once—and then only to disparage Kant's "architectonic"—Dewey's aim and conclusion are essentially the same. Cognitive activity and decision-making, as these bring new order into our lives, give rise to recapitulation, expression, celebration in some concrete objective and socially available form; such expression is the very substance of art, whether called by that name or some other.

What is missing from Kant and Dewey both—and what is surely not easy to supply—is the developmental aspect, the theoretical formulation of which allows us to say that the conditions required by the cycle have been met. For Kant, reason appears full-grown as a universal and necessary capacity of the human mind. His method is predominantly that of the analysis of knowledge or morality as something given in adult experience, not as something in process. Dewey's method more nearly follows the course of experience. But in spite of many important insights along the way, Dewey's method is too purely reflective and too little derived from empirical observation to provide even the outline of a theory of the "life of reason" from the beginnings of learning. But when we examine the subject thus suggested, we find that our sources are very restricted. Wertheimer's pioneer work[14] has unfortunately remained pioneer for the most part, and what we have for the study of growing rational capacities in children is cast in the developmentalist mold. I believe this mold is inadequate, though not thereby useless, to the cause of educational theory.

---

[14] Wertheimer, *Productive Thinking*, enl. ed. (New York: Harper & Bros., 1959).

A major fact of early education is its linkage, its interaction, with the postnatal stages of embryological development. Recognition of such linkage gives plausibility to a developmentalist view of education. The implicit metaphysics of this way of thinking is dominated by botanical or, rather, horticultural metaphors. Development is an unfolding from within, a cumulative realization of potentialities. There is no denial in this way of thought of external necessary conditions or of their "influence" on development. Children selectively assimilate the order of their environment, including the human culture. Such external materials of development are presumed to be there for the taking and if not, the child is deprived and his growth thwarted or stunted. The role of the teacher is to garden—to water, weed, spray, and fertilize, to cultivate. But in contrast with S-R theories the essential causality of the process is all internal. The rose may prosper or suffer because of the gardener's treatment, but it will be just that kind of rose, no matter what. So for education, the implied counsel is that of provision and remedy. Suitably provisioned, the process expresses an inner-directed schedule of developmental stages, each "laying the ground" for the next.

The legitimacy of such a descriptive metaphysics for education is relative to the range of questions we may seek to answer within it. In view of the manifestly inadequate provision for childhood in much of our society, the developmental viewpoint has implications we should preserve and insist upon. And where we can take for granted the prevalence of a sort of diffuse educational potential in a society—a material and human environment redundantly responsive to children's developing needs and abilities—this viewpoint is important by way of opposition to the nervous belief that everything learned or acquired must be "taught." For in such an environment children will predictably "grow up" without institutionalized planful intervention, without an educational technology trimmed to antecedently defined "behavioral objectives" of walking, talking, reading, carrying a glass of wine without spilling it, classifying, testing hypotheses, and so on. But when it is considered analytically as a mode of abstraction, the developmentalist scheme is inadequate to the deepest and most central concerns of education. It is inadequate because it buries under a metaphor just that level of interaction between "development" and "learning" without which our species would lack its most distinctive characteristics.[15]

What gets buried is the fact that learning is not only something which takes place within a framework which development provides, but that development is also something which takes place within a framework which learning provides. To make this reciprocity clear and emphatic, however, we must make explicit certain aspects of "learning" which schools have seldom recognized in practice and which educational theory has seldom insisted upon.

A great merit in the work of Piaget lies in the fact that he has directed his empirical research from the initial Kantian insight that learning is a selective process which depends not only upon the objective content available in any specific learning situation, but also upon the ensemble of resources *brought to* that situation by the learner. It is easy to suppose, in the interest of associationistic or behavioristic principles, that any such resources are simply the cumulative summation of past learning. Even on such an account, the center of gravity in any particular learning situation is, so to speak, on the inside. But as we can learn from a long philosophical tradition (in modern times most significantly from Hume and Kant), this explanation of our intellectual *modus operandi* is implausible in the extreme. The discrimination of similarity and difference is hardly an ability abstracted from similarities and differences which somehow forced themselves upon us before such discrimination was yet possible. The systematic organization of sensory experience in terms of space and time is not a mere abstraction *from* experiences which initially lacked that organization. Thus the succession of our perceptions does not add up to a perception *of* that succession nor do changes of locality as experienced automatically generate a grasp of the coexistence of localities in a comprehensive spatial framework. The recognition of causal relations or of genetic identity is likewise not "given" by empirical instances in which these characters are somehow forced upon us by our senses.

---

[15] See Dewey's careful discussion in *Education and Experience* (New York: Macmillan Co., 1938).

The major contribution of Kant to the discussion of these matters was not—as philosophers seem sometimes to believe—the introduction of the phrase "synthetic a priori," but an insight into the central role of hypothesis in learning and knowing, of the active or constructive nature of intellectual activity as contrasted with the passivity of the mind assumed in associationistic theories. It is on account of this constructive process through which the mind organizes its sensory information that Kant builds up his analysis of the mind-dependent order of space and time, the categories of the understanding based on the elements of Aristotelian logic, and the guiding role of reason. These are not taken from sensory experience but brought to it, a synthetic apparatus native to the mind.

But we, unlike Kant, must ask the genetic question. Does this *sort* of apparatus, regardless of the details of Kant's analysis, somehow appear full-blown in the mind of a child? Or is it there, latent, only to be *developed* by the cumulative passage of experience (but not altered by it)? Or is it the product of a more evolutionary process in which such intellectual powers evolve *as* devices, expedients borrowed or invented and reinforced on account of their demonstrated efficacy in ordering experience and extending it? The last named view is not Kant's, though passages can be found which would support it. I would like to develop this view at any rate as the culminating phase of the present paper.[16] Such a view requires no abandonment of the idea of native capacities. At any given moment of his career a child *already* manifests some such capacities in his behavior or, as we may put it, has a mind. Clearly also the subsequent evolution, although it must be deemed for each child unique at some level of scrutiny, will be characterized by strong family resemblances and at a certain remove of abstraction it is possible to talk about general "stages" of human development. From evidence gained by horizontal sampling across many individuals, it is possible to construct a composite longitudinal picture of the probable evolution of single individuals. Such use of the comparative method can prove very fruitful and *may* disclose patterns of development which will also be found universally in individual cases studied longitudinally. *May* disclose; on the other hand it is a familiar fact that the comparative method produces, through averaging, a sort of bland linearizing which *may* also, in the individual case, have a very slight correspondence with reality. What appears secondarily in such averaging, namely, fluctuation of pattern between individuals of the same age, or inconstancies and inconsistencies within single individuals at a given age—to both of which Piaget has given the label *décalages*—may be tell-tale signs of the essential individuality of development in a common world and cultural milieu.

---

[16] Piaget's answer is that experience is not only "sensory," but is also experience of the mind's own operational functioning *in relation* to the environment. Such self-representation derived from a growing operational competence is the *empirical* source of that logical and mathematical knowledge for which Kant could find no "empirical"—i.e., sensory—origin.

But what is most important for the present argument is that, as it stands, the conceptual scheme of developmentalism is one that can only very partially answer to the demands of *educational* theory. For such a scheme aims to be able to tell us what children will be like independently of their detailed human (and educational) careers. To say this is not to imply that developmentalists think "environmental influences" are unimportant. It is simply to say that the effect of unexamined differences in children's individual careers is of necessity treated as a "residual variance" from an average pathway of development which reflects at the same time an average educational milieu. The criticism is thus not of the *beliefs* of investigators working in this way, but of their conceptualization of the problem and the means of investigation which they adopt to match that conceptualization, *as relevant to education.*

And there is, I think, another important qualification. The criticism that has just been made is, so far at least, identical with one made frequently by some educational psychologists in the United States who think of intellectual development as a mere "summation of learning" and thus susceptible of acceleration through subjecting children to little training regimes, and from whose viewpoint I wish to dissociate myself. Piaget makes a very nice joke of this possibility by referring to it as "la question Américaine." As a nation we are, it seems, addicted to accelerators. But in rejecting this view Piaget sometimes appears to reject learning as a *primum mobile* of intellectual development, thus appearing to contradict his own views of the processes of assimilation, accommodation, and equilibration.[17] This last triad does, after all, explicate very well one meaning of learning, though it is not that of some learning theorists. Learning of a developmentally significant kind (and, I would add, of an educationally significant kind) is not the sort of thing which can be reduced to little training routines, little time-tabled packages with their prespecified objectives and tests of attainment. That these are without significant developmental consequences is not hard to believe. What we are or ought to be concerned with in the theory and practice of education is the kind of educational milieu which *does* affect the growth of children as persons in all of their competencies; thus, it must not be an a priori conclusion that developmental dimensions are unresponsive to educational provisioning, either quantitatively or qualitatively. Yet in detaching development *conceptually* from dependence upon the kind of provisioning which a society makes for its children, or which particular adults make for particular children, and which devoted professional teachers seek to provide, we are in fact declaring the irrelevance of such efforts to any except minor variations on a theme set out ahead of time. The developmentalist framework simply does not answer to the concrete epistemological situation of a teacher, who not only can contribute to the richness of children's learning, but also can stabilize and support those intrinsically rare occasions, in the lives of children as of adults, in which the discovery of new order in experience can bring culmination, reward, and growth in new dimensions.

If the powers of the human mind (and its limitations) are in some sense congenital, they are also in a complementary sense the product of a cultural evolution which is faster than biological evolution but slower than individual maturation. This means that in a not negligible degree every human being has the potentiality to contribute in his own style of growth and learning to a social process which is as yet very far from frozen in finality. In this sense, then, the content of individual development not only recapitulates the stabilized achievements of others, but also in its variance and regressions from the pattern of those achievements—if you will, in its décalages—adds to them or alters them.[18]

---

[17] For example, in his remarks on education in *Piaget Rediscovered: A Report on the Conference on Cognitive Studies* (Ithaca: Cornell University Press, 1964), pp. 113-17. On the other hand, more recently in *Science of Education and the Psychology of the Child*, trans. Derek Coltman (New York: Orion Press, 1970), Piaget shows a lucid, modest, and insightful recognition of the great distance between his own developmentalist schemes and the problems of education. One hopes that recent pedagogical schemes which apparently aim to "teach the stages" will be examined in the light of this eminent psychologist's own strongly reasoned disclaimers. But I take heart in my criticism of the Piaget "stages" by a passage in which he reasserts the basic developmentalist axiom, the existence of a [logical] "boundary between the contribution of the mind's structural maturation and that of the child's individual experience" (p. 172).

But the most important step in the argument is not yet explicit. For all its potential value, Piaget's specific theory of stages lacks in the present view an essential dynamical aspect. It is already implied, I think, in the previous discussion of equality, individuality, and freedom: when we view an individual history up to some particular time (and this is the epistemological perspective of a teacher) and project this against any standardized pattern of developmental stages, we *may not* legitimately use this projection as a mode of explanation, but only as a basis for questioning, as a device for gaining perspective. But in such perspective we will not see very much which is relevant to the art of teaching unless it is linked to other sorts of judgment which take account of individual characteristics of personal style and motivation. From such a point of view the décalages of development are not fluctuations from a mean (which has somehow mysteriously become normative), but suggestions as to further provisioning of the environment, further communications and guidance to help a child direct his *own* learning in accordance with his *own* resources and dispositions. In short, it is at this point that we must again give theoretical support to an idea implicit in all personal affairs and transactions—the idea of autonomy or agency.

And this brings us back to the role of reason. When a child or adult is operating in intellectually familiar territory, using categories and operations already well practiced, his activity belongs under the heading which Kant labeled as the *understanding*. But when his questioning leads him beyond intellectually safe ground (as from a *personally* secure base of operations it easily may), he is in a situation which calls for new expedients of thought. What before was new empirical content in the context of old thought patterns becomes rather a need to try on, to fashion new expedients of thought in a context of experience not fitting familiar rubrics.

---

[18] Compare Tony Kallet, "Away from Stages," *Mathematics Teaching*, Bulletin of the Association of Teachers of Mathematics (U.K.), 45, 1968. Republished in *Outlook* 1 (1970), Mountain View Center, University of Colorado, Boulder, Colo., 1971.

Such is the experience of a two-year-old friend of ours. After intense encounter with various musical instruments she first wanted to "find the music" inside the guitar—looking repeatedly inside to confirm, for us, the point of the question. Another day, she wanted to "see the music" inside of our phonograph. Consider also the question of an eight-year-old who had repeatedly harvested and planted spores from a mold to get new mold. He didn't want to know where *these* spores came from; he knew they came from molds. "But where do *spores* come from?" In both cases the questions outrun the resources available for answers, conceptually and empirically, so that some retreat to more modest (but related) investigations is in order. Yet in both cases a marker, an affective flag, has perhaps been placed beside a conceptual knot to be untied, and this knot will gather around it other later situations somehow alike, so that this first encounter (like the minute leader-stroke in lightning) will start a well-worn pathway in some future network of familiar relationships.

Such developments from the child mind to the adult mind, whether average in character or interestingly deviant, are neither the summation of many episodes of passive learning nor the fruit of some inner maturational process which owes nothing to conscious experience or self-conscious discrimination and choice. The two small illustrations were chosen to suggest the kind of situation in which a child's experience seems clearly to outreach his established intellectual capacities and cannot be dealt with either directly or by means of available transformations. Novelty is responded to, is vividly acknowledged, but accommodative learning is not yet able to provide a framework within which the novelty can be reduced to understanding. But this is the very situation which evokes what Kant calls the reflective judgment, which seeks to order and organize in the absence of assured determinative guidance. It calls forth that "play of the faculties" which may along the way evolve into a whole hierarchy of search procedures and even routines of inquiry by which we import new order into our lives.

It is this evolution (or the possibility of it) which meets the logical demand for the closure of the educational cycle—the evolution by which learning becomes a kind of career and evolves into the capacity to teach. It is the evolution for which I have chosen a traditional term—the life of reason. In encounters such as those suggested above, a teacher will not try irrelevantly to "explain," but instead will try to provide from many sources that which will enrich the empirical matrix of a child's thought or will suggest analogies and connections from previous experience which are, at that moment, unretrievable by the child himself. And although the outcome will never in principle be a foregone conclusion, respect for a teacher who has thus sometimes aided a child's growth is a source for that unique meaning of authority which Kant links both to persons as "ends-in-themselves" and, more abstractly, to the role of reason itself, which holds out in the face of novelty or disorientation the promise of higher levels of organization to be achieved.

To repeat, this account in no way implies that the processes of accommodative abstraction and learning are transparently visible to those who are the agents of these processes. On the contrary, the view of the participant is limited to that which he can at any moment bring to a focus in terms of his existing intellectual capacities. We can understand our own processes only, so to speak, in recapitulation, only from a point of view beyond those processes. And as teachers we understand the intellectual processes of children through an observant mapping of their behavior—a *reading* of their behavior—in terms of models of learning which we ourselves have built. (And which educational research might well both study as skillful teachers build them and then help clarify and formalize.) What we read is thus not imputed wholly to a child's "conscious self" but to the *child*, to the child as agent.

I have chosen these small examples deliberately from the sphere of natural and human phenomena as children encounter them to suggest that even in this domain the constructive, "poetic" process of choice—the experimentation with new ways of thinking and behaving—is the root of the dynamics of learning. In that sense the subject matter of Kant's second *Critique,* broadened beyond any exclusive reference to the specifically moral domain, is what is needed to complete a "developmental" theory of the mind. Choice in conduct is a search for ways of reorganizing conduct in the face of conflict, dissatisfaction, unbidden and forbidden impulses, new interests. Old habits of self-description and self-organization, already-made solutions, have been found wanting. In such situations the role of conjecture and improvization, the role of other persons as authorities, is not different from what it is in the domain we mislabel as "purely" intellectual. And again the aim of finality, of coherent and self-consistent organization, marks what we mean by "the role of reason."

But, then, as we choose differently so we grow differently, and as we grow differently we come along different paths to frequent congruence and communication with others. (If this is correct, a genetic ethics should undergird Piaget's "genetic epistemology"—the evolution of choice.) Conversely, for the lack of opportunity in the presence of stifling coercion or from anxieties which drain creative energies, we remain undeveloped and primitive in our resources, at best compensating for what we have missed by other developments which bring satisfaction and respect. Of such a kind, I believe, are some of the "residual variances," the décalages, which mark most human careers as measured against standard patterns of attainment. These latter can define, no doubt, a common framework within which the residue of our growth can be mapped, at least in part. But the dynamics of the process, the dynamics of learning and choice, is more open-textured, more discontinuous and incomplete, more redundant and more surprising in its novelties, more advanced and more regressive than horticultural analogies will allow. No more standardized account is answerable, I believe, to the perspectives of education.

CHAPTER XIII

Autonomy: A New Paradigm for Research

LEONARD I. KRIMERMAN

> *Science is more than the mere description of events as they occur. It is an attempt to discover order, to show that certain events stand in lawful relations to other events. . . . We cannot apply the methods of science to a subject matter that is assumed to move about capriciously. Science not only describes, it predicts. It deals not only with the past but with the future. Nor is prediction the last word: to the extent that relevant conditions can be altered, or otherwise controlled, the future can be controlled.* If we are to use the methods of science in the field of human affairs, we must assume that behavior is lawful and determined [My emphasis in Roman].
>
> B. F. Skinner, *Science and Human Behavior*, 1953.

The Alleged Incompatibility of Science and Human Freedom

Skinner's position, in varying formulations, is widely accepted throughout the social sciences.[1] Few social scientists see any need to *demonstrate* that science is inseparable from determinism and control, much as they see no need to prove that science involves verifiable hypotheses, or replicable and controlled experiments. The notion that a set of antecedent conditions and a "covering law"[2] can be located to explain and predict any human act is a common regulative or methodological principle. But is there a scientifically respectable alternative to, and way of restricting the scope of, the covering law model? Even those existentialist thinkers who find truth in the supposition that men are free and responsible agents often infer from this supposition that the methods of science cannot provide—or actually block—a complete account of human activity. We thus arrive at the *incompatibility dogma* (ID), according to which one can take seriously either free choice or scientific inquiry, but not both.

---

[1] See, for example, Leslie A. White's *The Science of Culture* (New York: Farrar, Straus & Co., 1949); Robert K. Merton's *Social Theory and Social Structure*, rev. & enl. ed. (New York: Free Press, 1968); and Robert R. Brown's *Explanation in Social Science* (Chicago: Aldine Publishing Co., 1963).

[2] The conception of a "covering law" and its role(s) within scientific inquiry are discussed in the introduction to Part II of my anthology, *The Nature and Scope of Social Science* (New York: Appleton-Century-Crofts, 1969). The classical expression of the covering-law analysis is presented in C. G. Hempel and Paul Oppenheim, "The Logic of Explanation," which first appeared in *Philosophy of Science* 15 (April, 1948).

In what follows, I shall argue against this received doctrine by advancing a positive program for studying human freedom which nonetheless adheres to the criteria for scientific objectivity. First, though, it is essential to look more closely at ID. On this view, to ascribe choice (voluntary or autonomous action) to a person is scientifically indefensible. For any such ascription presupposes the antideterminist assumption of *avoidability*—that people, under given sets of physical and sociocultural circumstances, can sometimes avoid doing *in the present* what they have previously done under those same circumstances, and that, without altering their circumstances, they could have *in the past* sometimes avoided what they in fact did. From the perspective of ID it is hardly surprising that social scientists like Carl Rogers, who would be loyal to both science and human autonomy, find themselves chained to a "contradiction . . . a deep and lasting paradox."

Underlying ID is a doctrine which makes the central business of scientific inquiry that of discovering general laws or uniformities by which whatever occurs (and thus all human behavior) can be explained, predicted, and controlled. Whether you want to produce men who believe that a proletarian revolution is inevitable or prefer a populace which rejects that inevitability, the behavioral scientist can provide a recipe for engineering the desired result. The laws sought, according to this doctrine, are all of a specific sort. They have the form, "Whenever A then B," and they affirm empirically discoverable but unavoidable connections between types of spatiotemporal phenomena. Thus, should an antecedent event occur (e.g., salt placed in water, glass heated, maternal rejection, rise in the price of a given good), this will suffice, given boundary and initial conditions, to produce a consequent event (salt dissolved, glass expanded, anxiety in the child, fall in demand for the good). The occurrence of the antecedent leaves one and only one kind of option. It eliminates all options (choices) that do not belong to the class denoted by the consequent term—the salt cannot but dissolve, the glass expand, the child become anxious, the good be less sought. To the extent that a choice remains between doing (bringing about) or not doing (not bringing about) the consequent, then to the same extent there is nothing unavoidable about the connection between it and the antecedent, and no scientifically reliable law has been discovered upon which to base predictions and explanations.

Moreover, for ID, the purported scope of what we may term "single-option" or "specific-response" laws is unlimited. Suppose we break down any given human act into its various elementary features —F^1 (involved anger), F^2 (took place in front of strangers), F^N, etc. Then, according to this view of science, for each of these features a covering law of the form "Whenever A then F^1," "Whenever B then F^2," etc., is waiting to be discovered. Given this set of laws, the act must be regarded as fully determined; there is no aspect or degree of choice about it. Each particular law makes a contribution to this complete unavoidability by eliminating every option outside a certain class (e.g., the operation of the first law prevents the agent from performing any nonangry action).

The notion of "boundary conditions" mentioned above deserves some comment. Salt dissolves in water, but only if it hasn't been stored in petroleum and if the water isn't already saturated with salt. The full form, then, of specific-response covering laws is: "Whenever A then B, unless C^1, C^2, . . . C^N," where C^1 through C^N indicate where the law is not expected to hold, i.e., place restrictions on the range of its projected application. There are many intricate questions concerning the logic and predictability of boundary conditions which cannot be pursued here. But one point, to be returned to when we later consider uniformities or laws that are compatible with human autonomy, should be made clear. To assure that a law expresses a genuinely unavoidable connection, boundary conditions which refer to human choices, preferences, decisions, and the like must be excluded. It would, for example, be unacceptable to proponents of ID to have our salt law state: "Salt will dissolve if placed in water unless men (or certain men) strongly prefer that it not do so, unless the water is salt-saturated, . . ." To permit this would be to concede that whether salt dissolves in water could be a function of human choice—that is, to concede that men might be able to avoid or prevent the occurrence of the consequent even when the antecedent has occurred, simply by preferring to do so. Such a "voluntaristic" boundary condition hardly seems compatible with specific-response covering laws, for these are designed to express unavoidable uniformities. For this reason we may refer to the laws crucial to ID's conception of science not only as "specific-response" and "single-option" uniformities, but as "antivoluntarist" laws. (I shall from here on use these three quoted terms interchangeably.)

Now, if specific-response laws are what every scientific inquiry relies on or seeks to discover, if their scope is unlimited, and if their boundary conditions are antivoluntaristic, then ID follows: Science can neither locate nor systematically study voluntary activity; indeed, it must deny that there is any such activity. To reject ID, we are required either to break the connection between science and single-option covering laws or to construct a fundamentally new conception of the laws scientists operate with and pursue. Nothing less will enable us to argue that human choice—in the full sense involving avoidability of actions men in fact perform—can be a proper subject of scientific inquiry.

But what of "statistical tendencies"? Many social scientists, it will be objected, do not presuppose determinism, nor do they seek specific-response covering laws. On the contrary, they work with and hunt for true statements like the following: (a) A middle-class Catholic American tends to vote Democratic in national elections, and (b) Nine out of ten college undergraduates who major in philosophy are moral relativists.

One reply here is that statistical tendency statements are commonly understood as approximations to universal laws. They are seen as a stage through which, regrettably, one must occasionally pass prior to locating unavoidable regularities. Nothing in principle prevents us from refining (a) into something of the form, "A middle-class Catholic American with feature F always votes Democratic in national elections." This change, for those who operate with statistical tendencies, would be viewed as an unequivocal improvement, since it would permit perfect, rather than imperfect, prediction and control. Discovering a true statistical tendency statement is merely an important first step towards discovering a specific-response covering law.

Secondly, what is the relationship between antecedent and consequent in statistical tendency statements? The occurrence of the former makes the latter, in varying degrees, probable. That is, *given the antecedent, it becomes unlikely that the agent could have done otherwise.* The set of antecedent conditions (being a philosophy major, being a college graduate) is thus a set of restraints or pressures which make choice difficult, and, in the preferred cases, virtually impossible.

At best, to focus exclusively on statistical tendency statements is to *ignore* human choice—to overlook, for example, the distinction between votes cast autonomously and votes elicited by external (cultural, physical) factors over which an individual did not have control. A man may vote, that is, *because* of beliefs and values into which he has been so thoroughly socialized that he is *unable* to imagine or to seriously consider alternatives to them. Even if the gathering of statistical tendencies does not strictly entail the denial of human choice, it certainly throws no light on how voluntary activity can be distinguished from nonvoluntary, on how autonomy or the capacity for choice is developed, on the factors that extend a man's range of options or that permit people to overcome obstacles (including social reinforcement) which block given types of choices. It seems fair to say that even if the hypotheses and conclusions advanced by statistical-tendency social scientists are not incompatible with human choice, the type of inquiry to which they restrict themselves prevents them from identifying and systematically studying voluntary activity. Thus, in its own fashion the statistical tendency framework also remains harnessed to the incompatibility dogma.

The Science of Autonomy and the Identification of Autonomous and Nonautonomous Actions

Social scientists, including educational researchers, thus typically accept some version or another of ID. It is my belief, however, that what can be called the *science of autonomy* furnishes a reconciliation of choice and scientific method, and thus stands in clear opposition to the contemporary social sciences of control. This new science possesses all of the essential marks of scientific objectivity—procedures that generate publicly testable observations, predictions, uniformities, and theories. But in addition it explicitly affirms that men can be and sometimes are free agents, that science can discover when they are and when they are not, and can further reveal how the ability to choose freely comes to life, how it can be maintained and increased, and by what it is threatened or diminished. Further, the science of autonomy permits the disciplined investigation of neglected issues arising from all domains of human life, including education. For many (apparently crucial) questions in educational research cannot be formulated or studied without presupposing decision, responsibility, avoidability. For example, (a) How do people acquire the capacity and desire to critically examine, discard, and find alternatives to beliefs and values into which they have been socialized? (b) Does an environment where the learner has prime responsibility for deciding the pace and content of his own education promote more or deeper learning than one with a preestablished curriculum and rate of progress? (c) How important for intellectual or moral or personal growth is a reliable context of trust and acceptance? Is such a context equally important for each kind of growth? (d) If children are free to choose what, how, and when they want to learn, does this increase or decrease their capacity to tolerate racial, ethnic, and religious diversity, to resist propaganda and attempted indoctrination, to identify and oppose, in themselves and others, the desires to dominate and be dominated? (e) Are freely adopted beliefs and values more readily exposed to criticism and less intractable than those not thus adopted?

The science of autonomy rests on three major distinctions: first, that between *autonomous* and *nonautonomous acts;* second, that between *enabling* and *sufficient conditions;* and last, that between *deterministic* and *antideterministic laws or uniformities.*

There is much to be learned from the works of third-force, humanistic, and existential psychologists. They stress that knowledge of man must be primarily concerned with distinctively human activities, e.g., those expressing guilt or self-consciousness, or those which are "fresh, spontaneous, interesting [and] show a new element in the agent's behavior."[3] And they work from normative or value-directed conceptions of human nature. When, they ask, is a man fully human; when can it be said that he is a whole person; that he is flourishing, well-functioning, that he gives a norm (itself to be transcended) to mankind, rather than merely exhibiting typical or normal patterns? What are the experiences, feelings, powers, crises, and achievements that make a human life worth living?

Still, it is not generally easy to be certain what thinkers like Maslow, Sartre, Rogers, May, and Buber would count as showing that a person is self-directed, possesses an integrated and authentically emerging self, or has participated in an I-thou encounter. Let us suppose that Smith is an atheist and a draft resister. Can anything be inferred from such data as to whether, or to what extent, he is self-actualized or a responsible maker of choices? The *content* of a man's values and opinions is no guide to whether he is their captive, to how freely he adopted them, or to his present capacity to assess and rethink them. If content is no indicator, how then do we confirm or disconfirm—make sense of—claims about responsibility, spontaneity, self-determination, becoming a full person? Without an answer to this sort of question, it will prove impossible to identify cases of the concepts central to the existential-humanistic view of man. What is taken as self-actualizing by one observer may be seen by others with equal sincerity, intelligence, and information as depersonalizing, alienating, inauthentic. The consequences of this conceptual indeterminacy are not merely conceptual. For if we cannot distinguish persons from objects, self-actualized individuals from those who lack any distinctive human excellences, then we can have no hope of finding reliable ways of liberating the former from within the latter, nor of developing (Rogerian) men "from whom creative products and creative living emerge." That is, the indeterminacy of our principal concepts will disable us from advancing any useful or even coherent account of the factors which prevent and restrict or promote and extend "initiative and originality," "the whole person," "dialogue between thou's." We will be unable to assist ourselves or others in enlarging the scope of choice and autonomy or in moving towards what we think of as the most distinctive and desirable forms of human activity.

---

[3] Rollo May, "The Emergence of Existential Psychology," in *Existential Psychology*, ed. Rollo May (New York: Random House, 1961), p. 36.

A primary task, then, for a science of autonomy is to provide indicators, or marks of evidence, by which disputes over whether a human act is autonomous or not can be settled. One way to begin is to distinguish three types of human conduct. The first would be purposive and voluntary (autonomous), the second would be purposive but not voluntary, and the last would be nonpurposive and nonvoluntary. For example, if Jones snores or sneezes, or if he starts bleeding from a wound, these reactions would normally fall under the last category. They are not purposive—a matter of what Jones wants and how he views his situation in the world—nor are they autonomous choices. Consider these examples: A person, acting under a posthypnotic suggestion, trying desperately to put out an imagined fire. A person unable to approach the edges of or look out from high places. An American marine into whom contempt and hatred for "gooks" has been indoctrinated by basic training. In the first case, there is no doubt that our individual *believes* the building is on fire; in the second, that he *fears* heights; in the third, that he *wants* to slaughter the Viet Cong and *believes* them to be subhuman in various respects. The behavior in each of these cases, then, expresses the agent's beliefs, feelings, and goals; in this sense it is purposive. Yet it is very doubtful that this behavior involves free choice or autonomy. The desires and convictions of the agent may not be ones he freely adopted or could now discard. Environmental pressures may have made them unavoidable.

In order for a human act to count as both purposive and voluntary it is crucial that its "intentional side" arises from choice. That is, not only must that act manifest goals and opinions (that of killing "gooks," that there is a fire), but the adoption and maintenance of those opinions and goals must represent voluntary activity. Thus, "When are a person's beliefs and desires autonomous?" is a question we must answer in order to distinguish purposive and voluntary acts from those that are purposive but not voluntary. Consider then the following preliminary account of the criteria for identifying cases of *autonomous belief* and *autonomous desire.*

1. *Autonomous belief.* The following considerations each lend separate support to the claim that a man's belief in a proposition, P, is autonomous:

   a. He understands P, in the sense that he can explain its sense to others, knows when it would be true and when false, and without help can use it and its word elements in circumstances (substantially) unlike those in which they were first introduced to him. Further, he can use P imaginatively, e.g., within the context of a story he makes up.
   b. He has in some way tested P and evaluated it against alternatives and has done so even though there are no extraneous rewards, such as status, income, or psychological serenity, for such testing and evaluation. Further, there are no (social or physiological) rewards, e.g., social approval, higher wages, more relaxation, dispensed to those who test and evaluate P, nor does he believe that such rewards are being dispensed.
   c. He does not change his (degree of) belief in P when extraneous rewards for believing it are either increased or decreased; nor does he when he thinks such rewards have been altered.
   d. He relinquishes or decreases his belief in P when relevant counterevidence is presented, and this occurs even though no extraneous rewards for such a shift are operative.
   e. When objections or alternatives to P are presented, this does not prevent him from functioning at his usual capacity, nor is he angered or threatened by cogent criticisms of P.
   f. When he is temporarily disabled from testing, asserting, arguing for, or acting upon P, this does not incapacitate him, nor does it prevent him from testing, etc., other beliefs.

2. *Autonomous desire.* The following considerations each lend separate support to the claim that a man's desire for a goal, G, is autonomous:

   a. He can state what attaining G consists in, distinguish it from other associated and similar goals and indicate how he would achieve it in circumstances (substantially) different from those in which G was first presented to him. Further, he can relate imaginatively to G, say by inventing a story in which a character pursues that objective.
   b. He has personally experimented with alternatives to G in cases where extraneous rewards were not dispensed to those carrying out such experimentation.
   c. When the extraneous rewards for pursuing or attaining G are either increased or decreased, this does not affect the degree to which he pursues it; e.g., if praise or approval for seeking G diminishes, this alone will not affect the probability of his attempting to realize it.
   d. He comes to curtail or eliminate his pursuit of G when certain of his autonomous beliefs concerning G and/or his relation to it alter; e.g., that he is capable of realizing G, that efforts to

attain G are compatible with most of the other aims he has adopted.

e. When he is exposed to criticisms of G as worthwhile, as satisfying, as achievable, or to objectives represented as preferable to G, this does not disable, threaten, or anger him.

f. When temporarily prevented from pursuing or attaining G, this does not incapacitate him, nor prevent him from pursuing or attaining other goals.

The relations between these criteria and autonomy are by no means simple. For example, none of them, by itself, entails that a belief or goal is autonomous. To make that judgment, all of the tests or factors must receive consideration. And in some cases, for example, 2e, failure to satisfy a criterion does not suffice to establish that autonomy is lacking. A useful approach to these criteria, suggested to me by D. Bob Gowin, conceives them as a cluster of factors specifying an *ideal type*. Where all twelve factors obtain, we have our strongest or ideal case and can speak of an act as unequivocally or paradigmatically autonomous. If nine or ten are present, then the act should be judged autonomous, but not without some hesitation and qualification. Where an agent fails to satisfy six or seven of the criteria, it becomes misleading to either assert or deny that he acted autonomously. We are faced with a borderline case, about which no clear decision can be reached. If none of the twelve factors obtain, any claim of autonomy can be decisively rejected. (We would have an ideal case of nonautonomous behavior.)

Despite this complexity, criteria such as those proposed can assist us in deciding whether or not to count specific human acts as autonomous. If the act is not purposive, does not manifest the agent's desires and beliefs, then it is not autonomous. If it expresses his views and objectives, but these do not satisfy *any* of our criteria, then once again it is safe to classify the act as nonautonomous. If the agent's desires and beliefs satisfy all of those criteria, we can then judge his act to be unequivocally or ideally autonomous. If they satisfy almost all the tests, the act should be regarded as autonomous to a high though not maximal degree. Where there is considerable criteria conflict and a roughly equal number of criterial factors are present and absent, this will establish that (and why) no decisive case for affirming or denying autonomy can be advanced. Such situations are analogous to the case of a ball player who hits thirty-five home runs a season, but whose batting average is never higher than .215. He is not clearly a bad hitter nor clearly a good one. He cannot be accurately classified as falling at either extreme of the continuum running from good to bad hitters, nor for that matter can we simply assign him an intermediate position between extremes. In the same way, many of our actions will neither be clearly autonomous nor clearly lacking in autonomy. One merit of the ideal type analysis presented above is that it provides ways of identifying such borderline cases and of objectively distinguishing them from those acts about which unequivocal (or clear though qualified) decisions as to autonomy or its absence can be made.

The two sets of criteria are designed to dispel some of the conceptual haze that often obscures the writings of those who maintain that what is distinctive about human activity cannot be reduced to "simpler" physical-chemical-biological reactions, or is not accessible to the methods of natural science, or indeed to any form of scientific inquiry, if science is based on the ideal of control. The criteria specify or are logically connected to observable performances, e.g., distinguishing one goal from another, inventing a story, testing a proposition, and so forth. The recognition of those performances should be as objective, teachable, and verifiable as the recognition of the behavioral criteria for being a strict parent, an accomplished violinist, or a skillful juggler.

Why *these* criteria? In part, because they exclude certain major forms of involuntary (unavoidable) conduct such as compulsive or phobic behavior and behavior which is under the control of environmental pressures. That is, Jones may believe whatever his boss or priest asserts, or whatever is assigned prime space and time by the mass media. His acts may have a more physiological basis—painful prior experience may make it the case that when he is presented with the color red or a certain odor of decay, he becomes apprehensive and is thus led to reject whatever goals and beliefs are then suggested. It may be that he will predictably believe or pursue whatever is proposed by young women with chalk-white skin and jet-black hair, since this combination above all others has in the past excited him sexually. Such cases of environmental programming do not manifest choice. It is an advantage of our analysis, given criteria 1b-1d and 2b-2d, that they cannot be counted as autonomous activity.

Alternatively, Jones may be a compulsive gambler, a cigarette addict, or phobic concerning birds or loud noises. He may be unable to credit anything a person says once it becomes evident that that person is homosexual. There may be no explanation for such conduct in terms of social or even physiological reinforcement. Nevertheless, Jones may still be incapable of controlling his propensity to indulge in it. The last three tests in each set are intended to help separate behavior arising from compulsive, addictive, or phobic beliefs and desires from the class of autonomous actions.

The Science of Autonomy and Explanations by Reference to Enabling Conditions and Antideterministic Laws

Hopefully it is now possible to reliably distinguish autonomous from nonautonomous acts, voluntary activity from behavior which is not freely chosen. However, the science of autonomy does not stop with classification, though this is an essential first step. The impulse behind such a science is to discover ways of extending the scope of human freedom, to isolate factors which foster the capacity of human agents to voluntarily adopt courses of conduct, values, beliefs, and so forth. Central to every science is the search for explanation, the attempt to locate "causal factors," i.e., those which give rise to, influence, produce, predispose, shape, underlie, etc., other phenomena. But can there be "causal factors" for an autonomous act? Will not a science of autonomy entirely reject explanations?

There is, however, no reason for exiling *all* forms of causal explanation from the new science, though, admittedly, some forms are incompatible with its focus on human autonomy. Thus, any act which is shown to be autonomous by our previously stated criteria cannot itself be brought under—predicted or explained by—specific-response regularities of the form, "Whenever A then B." For if the act was an occurrence of type B, then the occurrence of some event of type A rendered it *unavoidable*, and thus not a matter of choice or of autonomous activity. To put this another way

—to establish a lawlike explanation for what a man does is to locate circumstances and uniformities which are *sufficient* to produce his behavior. To explain why the dog barked in terms of a sound having a certain frequency being produced under given conditions is to imply that dogs bark whenever that sound is produced under those same conditions. The beasts' barking does not arise from choice but is predictable from conditions external to them.

If Jud voted for Wallace autonomously, there is nothing we can discover that is sufficient to produce acts of that sort, nothing such that whenever it occurs, a person will invariably vote (autonomously!) for Wallace. Should we come upon some single-option covering law or sufficient-condition explanation, this will show that Jud would have voted as he did, *regardless of what he desired or believed* (or that he was unable to alter his desires and/or beliefs). But this means that his behavior was *unavoidable*—the vote may have been purposive, but it was not autonomous. Now Jud may of course have voted for Wallace because of environmental conditioning or because of goals and opinions to which he is attached addictively. In this sort of case an explanation, in terms of phenomena sufficient to elicit his vote, is certainly possible. Once it is established, however, that the act to be explained is autonomous, the search for this kind of explanation becomes futile.[4]

To exclude sufficient-condition explanations of autonomous acts does not, however, bar all types of causal or scientific explanation from the science of autonomy. If a congressman supported a bill providing for low-cost abortion on demand, after a history of opposition to such measures, we would certainly seek an explanation. And we would do so, even if the act were credited as sincere and as voluntary. The same is true of a child who "somehow" finally learns to judge pop flies, or to curb his fear of darkness, or to stop enjoying the role of bully. In such cases, we are not satisfied with merely "understanding" the agent's beliefs and preferences, even if these have recently altered and even if they do provide insight into why it was rational for the act to be performed. More is sought. For example, *something* of an explanation may be furnished by the fact that the has-been bully came to develop a sense of fairness according to which no one should interfere with anyone else, save in self-defense. This should, however, strike us as incomplete. How did this sense come to life, especially since it had enjoyed no previous access to the child's personality? Why was the congressman convinced, after so many years, that he could or should advocate the abolition of all laws prohibiting abortion?

---

[4] For a sophisticated defense of the view, which I here reject, that voluntary acts *can* be explained by causal uniformities of covering laws, see Wilfrid Sellars's "Fatalism and Determinism," in *Freedom and Determinism*, ed. Keith Lehrer (New York: Random House, 1966). In the same anthology, positions similar to mine are presented by Chisholm, Danto, and Ginet.

The acts for which we here seek explanation, let us assume, are voluntary. The child's development of a new skill or attitude and the congressman's altered vote are instances of neither conditioned nor compulsive behavior. Sufficient-condition explanations by reference to antivoluntaristic covering laws can thus be ruled out. In what way(s) *can* these acts be explained?

It is at this point that the concept of an enabling condition will prove useful. Consider the following cases:

1. Congressman Andrews has always supported laws prohibiting abortion out of fear of losing his seat in a largely Catholic district. In the past few months, however, he has noticed several changes in his constituency: (a) overpopulation and Women's Liberation have become popular causes with younger college-educated Catholics; (b) these Catholics see abortion as they see birth control—matters of personal conscience, not legitimately regulated by government; (c) many old-time Catholic families have left the district because of an influx of migrant workers who have no antipathy towards legalized abortion. With an eye on these changes, the congressman votes to legalize abortion.

2. Congressman Andrews goes to a dinner and fund-raising event which younger party-workers unexpectedly turn into a full-scale sensitivity marathon. No one addresses him in the superficially polite and ceremonial manner he has come to expect from ordinary citizens. People touch him, curse him, comfort him, who were total strangers moments before. They are willing to share with him feelings and experiences that usually remain unmentioned and well guarded. He finds himself drawn out into the open, trusting others and being trusted. Several women confide to the group that they

have had illegal abortions; one of these accuses Andrews of contributing by his votes to the death and damage produced by this practice. Why should any group, she demands to know, be entitled to force their religious views concerning the "rights" of the fetus on others of opposed convictions? Whose rights—those of mothers constrained to years of housework and child care for unwanted youngsters or those of unborn parasitical organisms—should take precedence? Andrews concedes that he has no answers to such questions and that he has rarely considered them. For him, what has always mattered most is which side of an issue will gather the greatest number of votes. This admission evokes two sorts of responses. One group repeats the earlier accusation, labeling Andrews a hireling and executioner for the Roman Church. Another criticizes the congressman's relationship to his political office: Why is he so chained to it, why does he depend upon it, much as a teen clings to childhood to secure an excuse for not standing on his own? What is he afraid might happen if his own judgment and personality came to life and began to guide his conduct?

Here Andrews becomes angry with the group. Walking from person to person he asks each one how long and how well they know him. He then claims that everyone is the slave of *some* set of goals or attitudes. The group is therefore doubly dishonest, since they are condemning him on paltry and third-hand evidence for what they themselves are also guilty of. As for the blood of illegal abortions, why didn't the mothers present examine their own hands? There *are* foundling homes and adoption services. Once again, they have attacked him, and him alone, where they are equally vulnerable. As he turns to leave, a woman's voice sings out, "Congressman, did that last speech win or lose votes?" Andrews begins to answer "Who knows?" and then suddenly realizes that for a time anyway, he hasn't given a damn about votes, about what others expect or want from him, about winning a group's support in order to vote in ways that keep their support. Surrounded by much warm laughter, he breaks down, partially from joy, partially from a sense of potential shipwreck, partially from exhaustion. Several days later, he publicly pledges to help repeal the current antiabortion law.

In both of these cases, we could say that Andrews had been

E^2 was thus of use in countervailing several distinct obstacles; for this reason, it is to be ranked higher than E^1 in the set of enabling conditions for the congressman's act.

Suppose we want students to learn not only the external facts about a group of poems, but to master each poet's way of responding to the world, to his own feelings, and to his medium. Three methods are proposed. The first has the instructor lecturing during each class period and setting short-answer or brief essay examinations throughout the semester. The second is wholly unstructured. It dispenses with grades and exams and determines what is studied and how it will be discussed from period to period by group vote, though no one is really bound by these decisions. The last relies on a planned structure, but one that is individualized. Each member of the group, on the basis of his own current attitudes, is to select certain poems each week and to live (think, feel, perceive) by their aid as fully as possible. He keeps a journal of his efforts in which he evaluates his grasp of the poems and how they have altered his life. This journal is critically responded to by an advisor and periodically shared with the class. A comparison of these three approaches can now be made to determine the extent to which they serve to countervail obstacles to the sensitivity we hope to develop. Quite conceivably, one of the three may outrank the others in this regard. The individualized-structure approach might be found to overcome student passivity and, in addition, to overcome the alienation often felt towards material that is overly academic or "irrelevant"; this may not be true of the others.

The point of ranking enabling conditions is to augment the kinds of explanation open to the science of autonomy. Thus far this science has been limited to explaining the occurrence of specific autonomous acts by appealing to (unranked) enabling conditions and the obstacles they countervail. But in most or perhaps all cases enabling conditions for acts other than the one which actually took place will also have been present. Why, then, was that act performed, rather than any of the alternatives?

To appeal to unranked enabling conditions *can* serve to indicate in what ways a given autonomous act, A, had been prevented previously and why it was not thus prevented in the case at hand. However, it cannot help resolve the further issue which is bound to arise: Why was A chosen, as opposed to other acts, obstacles to which also were countervailed? Why did Andrews come to sponsor legalized abortion, rather than either refuse to take any stand, or oppose legalization but improve the conditions in, and accessibility of, foundling homes? Suppose, though, that we compare the various sets of enabling conditions and the rankings of their members and find that high-ranking enablers were present for act A, but not for other autonomous options. This finding will go towards explaining what unranked enablers cannot explain, i.e., why A was selected over those other options.

It appears, then, that the question "Why did A occur?", which seeks an explanation for an autonomous act, can be given at least two interpretations. First: Given that some obstacle has blocked the performance of A in the past, what is that obstacle, and by what was it overcome in the present case? The second sense of "explain an autonomous act" does not focus so narrowly on the act itself. Here the demand for causal explanation can be reexpressed as follows: Why was A performed, rather than B or C, when enabling conditions for all three were available? When this second sense is involved, an appeal to unranked enabling conditions will not suffice. What is needed in addition is an account of the power or scope of the enablers that obtained for A, relative to the power or scope of those that obtained for B and for C.

The ranking of enabling conditions, where justified by replicated studies, can generate not only explanation in a distinct sense of "explanation," but also predictions and even a certain kind of uniformity or law as well. Let us imagine that E^1, E^2, and E^3 are enabling conditions for act A and that E^4, E^5, and E^6 serve the same function for B. All six conditions, we know, are elements of a situation that will occur in the future. Can we predict whether A or B will be chosen? Let us suppose that the conditions have been ranked and that, whereas E^4-E^6 are all single-function enablers, E^2 countervails several distinct obstacles. With this data it would begin to be reasonable to predict the performance of A. Or, more carefully—as with any scientific uniformity—we could predict that other factors presumed of equal weight, e.g., the agent's desire to do A versus his desire to do B, A will be selected.

Clearly, the prediction here is based on—makes an application *enabled* to vote for legalized abortion. In the first, certain facts about his constituency enable him to discard his long-standing fear that by siding with the legalizers, he will lose his seat. The explanation in the second case might run as follows. It was the congressman's inability to act or think independently and to take moral considerations seriously which had prevented him from holding any position not tolerated by the majority of his electorate. Through the openness, emotional intensity, and unrestrained encounter of the marathon, these inabilities came to be recognized and (partially) overcome. He was thus enabled to move away from his previous stand on abortion.

The explanatory model here involves three elements. First, *an autonomous act, A,* the explanandum. Second, *an obstacle, O,* which has deterred or prevented A, or acts of the same type, from occurring. And last, *an enabling condition, E,* by which O is overcome, in this sense: With E, anyone can, if he so desires, perform A. Or, more precisely, anyone can, if he so desires, perform A as readily in the presence of both E and O as in the absence of both. It is important to note that E is neither a necessary nor a sufficient condition for A. E's occurrence need not and will not invariably be followed by acts of that sort. Further, in the absence of E, other enablers may serve to overcome O, or O may itself never arise to block A. The distinction between enabling and, in particular, sufficient conditions is essential to the science of autonomy and to its explanatory model. For if E is, or is part of, an empirically sufficient condition for A, then an antivoluntarist regularity obtains, and our agent's autonomy is a fiction (see page 339 above).

There are at least two ways to safeguard the distinction. The first is to insist that E cannot provide an enabling condition for A unless cases can be found in which E occurs, but the agent or someone very similar to him does *not* perform A. Thus, in case (2), we would need to inquire whether other congressmen had retained their antilegalization positions even though they were subjected to highly charged encounter sessions at which those positions came under attack, where trust was given and accepted, and so on. Another way of preserving the distinction is to require that for E to count as the enabling condition which explains A, some other condition, F, must also have been present, where F is an enabling condition for another act, B, which is a contrary to A, in the sense that no one can perform A and B concurrently. In case (2) F might be a three-hour session with Roman Catholic clergy, by means of which the beliefs and attitudes initiated by the marathon are sharply criticized and brought under reconsideration.

It may well be that no enabling-condition explanation should be thought of as complete unless it invokes *both* of these ways of blocking the entry of sufficient conditions. In any case, where both are established, this would plainly help to ensure, first, that the supposed enabler does not generate some new conditioning pattern or function within a more fundamental single-option covering-law explanation to preclude alternative choices, and thus, second, that the act being explained can be viewed with warrant as autonomous.

In brief, one way to bring explanation consistently into the science of autonomy is by identifying obstacles to autonomous acts and the conditions which, while enabling agents to overcome those obstacles, do not themselves elicit or suffice to produce the acts in question. Much more, however, can be said here. Thus far, through enabling conditions we can gain some understanding of the causal factors for specific and isolated past actions. But what has been advanced cannot yield much generality, nor can it be directly employed in the service of *extending* human autonomy. One further step to be taken is the *ranking of enabling conditions*. Thus, given a certain type of (autonomous) human action, it is possible to construct a list of obstacles—empirically and conceptually discovered—to its performance. Certain of the act's enabling conditions will be useless in countervailing all but one of the listed obstacles. Others will be increasingly good resources, until we reach enabling conditions that can be employed to overcome most or even all known obstacles. An illustration of this range is provided by cases (1) and (2). In the first, certain fresh information (E^1) enables Andrews to eliminate his fear of losing office should he vote for legalization. It seems evident that this information is a fairly restricted, if not single-function, resource. It serves to overcome the obstacle presented by "fear of losing the election by losing the Catholic vote," but hardly anything else. On the other hand, the encounter session (E^2) enabled our congressman to recognize two separate inabilities, and, moreover, to start acting in ways they had previously deterred.

of—the regularity that whenever E^1-E^6 are the only enablers present for either A or B, the former will be preferred to the latter. Thus, the ranking of enabling conditions provides us not only with predictions, but with predictions made in accordance with general laws.

Two questions on the ranking of enablers deserve attention before any further exploration into the science of autonomy is attempted. First, are these rankings established objectively and verifiably? Second, how can enabling-condition covering laws (EC laws) be fitted compatibly into a science which represents itself as studying autonomous activity and is pledged to extend, rather than curtail, human choice?

In regard to the first question, consider the three groups of students discussed on page 344 above. Our first task, after providing acceptable indicators for the hoped-for sensitivity (S), would be to discover the obstacles to the development of S. These would no doubt comprise an open-ended set, but no more so than the boundary and standard conditions presupposed generally by scientific laws. Ordinary experience, supplemented by scientific imagination, would suggest hypotheses as to obstacles and types of obstacles—e.g., those involving teacher-student roles and relationships; those arising from student attitudes towards education, towards themselves and towards the opportunities they see held out for them; those peculiar to the material under discussion. These hypotheses would all be of the form: Given (presumed) obstacle O a person will be unable to either (a) autonomously want to develop S, or (2) develop S, though he wants autonomously to do so. By applying Mill's methods, particularly the *joint method of agreement and disagreement* and the *method of concomitant variation*, controlled observations could be made of the relation between O and the inability it is alleged to produce. In this way, the claim of any presumptive obstacle could be either confirmed or disconfirmed.

The same should also hold for the identification of enabling conditions. Hypotheses would again be generated, this time as to plausible (types of) enablers. They would take the following form: Given the presence of (presumed enabling condition) E, and of obstacle, O, to the performance of act A, anyone who wants to perform A will be as able to do so as someone who wants to perform A in the absence of both O and E. Once more, controlled observation, verification, and disverification would seem to be within reach. Holding other factors constant, do individuals who confront both O and E find it more difficult to perform A than those who operate in the absence of both E and O? If (and only if) they do not, can E be counted as an enabler for A.

Now if hypotheses concerning the obstacles to, and the enabling conditions for, autonomous actions can be put to controlled testing, then it must be possible to rank enablers objectively. For the scope or power of an enabling condition is merely a function of how many obstacles it can be employed to overcome, relative to other enablers, for the same piece (or type) of conduct. What then of our second question: If the ranking of enabling conditions provides law-based predictions, how can such conditions be compatible with, much less essential to, the science of autonomy? Are not EC laws of the form, "Given only this set of enablers for A and only this set of enablers for B, A will be performed rather than B," ways of saying that certain options are inaccessible and others unavoidable?

In reply here, it is essential to keep a sharp distinction between EC laws and those uniformities which involve unavoidable connections and which we have labeled "single-option," "specific-response," and "antivoluntarist" laws. One major contrast concerns the boundary conditions for the two types of laws. Single-option laws are deterministic, partly because they do not leave room for voluntaristic boundary conditions. But EC laws are designed to explain, and thus must remain compatible with, autonomous activity. Hence, they can include, as limitations on their scope, factors such as (a) the degree to which the agent autonomously wants to perform the various options in question and (b) the extent to which he recognizes and wants to make use of whatever enabling conditions obtain. Nothing in any EC law compels an agent to want to perform one autonomous act more than another; and nothing in such laws makes it unavoidable for him to be averse to the use of some or all of the enabling conditions for a given action. In brief, the boundary conditions for an EC law (as opposed to those for single-factor laws) *presuppose choice*, and they allow for the possibility of autonomously acting in precisely the way predicted against by the body of the law. *By our own choice, we can step* *outside those boundary conditions, thereby creating a situation in which the law is inapplicable.* When we do so, we do not contradict or falsify the law. Thus, since EC laws as conceived here are not falsifiable by men choosing to refrain from the actions they predict, they are compatible with human autonomy and place no restriction on the scope of our freedom. They render no option unavoidable and none inaccessible.

What we have called EC laws belong to a category of considerable importance for the science of autonomy, viz., *antideterministic laws.* These regularities permit prediction and explanation, but they are not specific-response uniformities. On the contrary, they are just those uniformities which presuppose human choice and autonomy. We have just seen how EC laws fit into this category; that is, by virtue of their boundary conditions including such references as "unless the agent autonomously prefers another option" and "unless the agent autonomously prefers not to make use of relevant enablers." There are, however, many other types of antideterministic regularities. Let us suppose that what we wish to explain is not why one act was selected over others, but *why a certain set or range of options was transcended.* All men have "limited repertoires" when facing some, and probably most, circumstances. In such cases, what we can attempt, or even intend, to do is fixed within narrow limits. An insult brings either a return in kind or a ruffled walkaway. The government's words are taken as either high-sounding but deceitful or as blunt and illegitimate appeals to force. Now it is a fact that men sometimes overcome such limitations and thereby find themselves or others possessing an extended range of options.

Thus, for example, when confronted by the absence of his mother, Dennis may invariably either throw a tantrum or withdraw into silence. Through trial and error, we may discover that should his mother leave after he has begun work on a project which absorbs him, and which requires creative exertion, he remains somewhat restless, but neither flares into temper nor retreats into noncommunication. Tentatively, we could then say that the kind of creative absorption he experienced (CA) is a "range-extending factor" for transcending his previously restricted repertoire. What holds for Dennis, however, might also hold for all kids, or for all kids of a determinate variety. In that case, a universal law would be established of the form, "Whenever a child (of type T) experiences CA, he transcends the fixed range of either throwing a tantrum or becoming withdrawn in response to his mother's absence." This law, which certainly appears to be empirically testable, would provide an indefinite number of predictions and explanations; nonetheless, it not only is compatible with autonomy, but requires or implies it.

A white person may display a limited range of responses in his relations with black people. Either he patronizes them or indulges in self-condemnation and expressions of guilt. Here, a range-extending factor (REF) might be a situation in which the white is assigned and plays out the role of a black constantly humiliated by well-intentioned whites (played by blacks). Furthermore, by Millian methods of controlled observation, we could test the hypothesis that this sort of role exchange is causally linked with the development, in similarly circumscribed whites, of a capacity to choose options beyond patronization and self-blame. If the tests yielded confirmation, it would be legitimate to claim that an antideterministic law had been discovered. Why so? Because the consequent, first of all, is a capacity rather than any specific action. And secondly, it can be further shown that the REF does not make a new-type of response (here, to blacks) unavoidable: the previously open options are still within the range of the agent. Thus, REFs do not diminish or destroy choice; rather, they expand its scope.

In *The Lives of Children,* which is a rich source for the science of autonomy, George Dennison writes:

> . . . since José early came to trust me, I was able to do something which, simple as it may sound, was of the utmost importance: I made the real, the deeper base of our relationship a matter of physical contact. I could put my arm around his shoulders, or hold his arm, or sit close to him so that our bodies touched, or lean over the page so that our heads almost touched. . . . The importance of this contact to a child experiencing problems with reading can hardly be overestimated. José was frequently at some sort of borderline in his sense of himself: struggling to comprehend, or to cope with his shame and frustration, or to cope with the excitement of little flashes of understanding. Little eddies of anxiety moved around and around him—yet he was not thrown back, in his need for support, upon attitudes or emotions; our physical contact gave him security. Nor was it my body beside his that meant so much,

> but the fact that my body vivified his awareness of his. He knew where he was: he was in his skin; and when the little bursts of panic made his head swim or his eyes turn glassy, he did not have to run away or reject the task *in toto*. He would gather himself together, because his real base—his body—was still there. I know that in the course of our lessons I committed errors and God knows how many omissions, yet this physical base was so important and so reliable that it provided all kinds of leeway. It took the sting (though not the seriousness) out of my rebukes, it expressed a concern I could not have put into words. . . . If one single formula were capable of curing the ills of our present methods of education, it would be this physical formula: bring the bodies back.[5]

Physical contact with a trusted adult here serves for José as a range-extending factor. When faced with the difficulties of learning to read English, he can only run from the lesson or remain but refuse to attempt any further reading. But when Dennison touches José or presses his body against that of José, the boy's scattered and anxious self receives a definite location, stability and unity return, and he transcends the prior limitation. So might it be for any child, like José in specific respects, who finds reading, or reading in a second language, hard to master; perhaps even for any child struggling and failing to learn any sort of material.

Thus, in addition to EC laws, the class of antideterministic uniformities also includes range-extending (RE) laws. The form of these would be, "Whenever a person (of such-and-such type) experiences R, he transcends a certain circumscribed range of options in response to situations of type S." Any such law has three main constituents: (a) R, a certain REF; (b) a certain kind of restricted repertoire, elicited by (c) S, a certain type of situation. RE laws, as mentioned above, do *not* predict or explain specific responses; their consequents can be viewed, rather, as capacities. And they do not rule out or eliminate options, but increase the number or variety of those an agent can choose to perform.

Furthermore, given these three sorts of elements in RE laws, it would seem that at least two types of *RE theories* can be formulated, where an RE theory is a universal regularity which explains or guides the search for lower-order RE laws. Thus, we can hypothesize about:

---

[5] George Dennison, *The Lives of Children* (New York: Random House, 1970), p. 169.

1. What kinds of REFs would work to transcend given types of restricted repertoires, regardless of the situations by which the latter were elicited?
2. What kinds of REFs would contervail given types of eliciting situations, regardless of what, in terms of fixed repertoires, they elicited?

For example, in response to (1), we could form a class of "dominator-dominated" restricted repertoires which would include cases where a child only related to other children as a bully or as a target for bullies and cases where a man was attracted only to women he controlled or to those who controlled him. We would then speculate that a certain type of factor, R, would work to transcend this *sort* of behavioral repertoire, regardless of the situations by which those behaviors were elicited and irrespective of how this fixed range might be manifested. This second-order hypothesis would yield the following guide for inquiry: Given any instance of the dominator-dominated fixed repertoire, look for an REF of type R to develop the capacity for transcendence. This sort of hypothesis functions to direct our efforts to locate more concrete RE laws, i.e., regularities between more specifically delineated REFs and restricted repertoires.

Again, we might classify children who either blindly guess at answers or keep entirely quiet together with college students who restrict themselves to either cramming for, or cheating on, exams. Perhaps John Holt would see both groups as exhibiting the "right-answer syndrome"—the learner here identifies learning with being able to recite right answers and strives to exhibit (or conceal that he lacks) this ability. Hypotheses could then be invented as to what type of REF could be relied on to break through the right-answer syndrome in any and all of its many forms. Conceivably, one REF here would be the substitution of a cooperative for a competitive learning environment. Our second-order hypothesis or theory would then advise: If you want to discover laws concerning the transcendence of specific manifestations of the right-answer syndrome, look for REFs involving increased cooperation in the learning environment. *This advice does not predict any specific REF, but it does tell us in what general area to look.* We may find that removing grades may work for one sort of person hung up on right answers, but not for another. Initiating the idea of having everyone work jointly to solve mathematics problems or to turn out a publication may succeed for only certain types of right-answer fixed repertoires. The theoretical hypothesis, in short, does not generate any specific and directly testable lower-order regularities. Nonetheless, under its tutelage we may be led to discover an indefinite number of such regularities.

In much the same way, research-guiding hypotheses can be devised in response to question (2) on page 351. Moreover, though there is not room to discuss them in detail, it is possible to identify several additional types of antideterministic uniformities. For example: (a) *first-person incorrigibility laws:* What factors enable a person to become incorrigible, in the sense that his sincere—and autonomous—first-person testimony is the most decisive evidence concerning his own feelings, desires, sensations, etc., rather than any appeal to "objectively measurable" physiological states or to his nonpurposive or nonvoluntary behavior? (b) *rational grounds laws:* Men whose beliefs are predictable from their hopes and ambitions are less autonomous that those who can form judgments on rational grounds which conflict with what they have a stake in; how, then, is the ability to form such unbiased judgments enabled? (c) *laws concerning sequential stages of autonomy development:* Are there pairs (triples, etc.) of types of autonomy such that a person can develop one type only if he has already developed a second, but not vice versa? Or, alternatively, are there certain kinds of fixed repertoires (e.g., the right-answer syndrome) such that only if they are countervailed by appropriate REFs can other fixed repertoires (e.g., those involving the inability to take stands on moral issues) be undermined effectively?

Summary and Implications for Educational Research

This compressed account of the science of autonomy needs a summary. The thesis has been that the autonomy paradigm, involving observable criteria for distinguishing autonomous from nonautonomous activity as well as explanations and predictions based on enabling conditions and on antideterministic laws of various sorts, points out a new direction for the scientist concerned with human conduct. Specifically, it provides a way to discard the incompatibility dogma; Despite focusing on and accepting the reality of free choice, the science of autonomy satisfies every essential canon of scientific objectivity. Its distinctive categories are teachable and constrained by activity-linked indicators (see pages 334-37 above). The new science can furnish (causal) explanations and predictions of autonomous acts by means of hypotheses testable through Mill's methods and by controlled observation (see pages 340-47 above). It will also generate uniformities on at least two distinct levels of abstractness, i.e., laws and regulative principles for locating laws (see pages 348-52). Hence, the *conceptual equipment* of the new paradigm renders autonomy as accessible to public observation as many less controversial human capacities and talents, e.g., testing a proposition, relinquishing pursuit of a goal, inventing a story (see page 337 above). Moreover, its *explanatory framework* of enabling conditions, EC and RE laws and theories is not only entirely compatible with, but pays special attention to, free choice (see pages 341-44 and 347-48 above). With the science of autonomy, one can take both scientific objectivity and human autonomy seriously; within it, one utilizes the former to gain insight into, and to help extend, the latter.

What are the implications of the science of autonomy for educational research? First, by undermining ID, the new science permits certain types of educational research to be initiated which, simply because they presuppose vountary activity, are now either neglected or deemed suspect of being "unscientific." For example, the science of autonomy can objectively investigate questions like those stated on page 332 above. Thus, instead of it being unscientific to be concerned with issues which recognize human choice and the development of autonomy, it should be seen as unscientific to ignore or reject them in favor of issues focusing exclusively on nonautonomous human behavior.

The second implication of the science of autonomy results from a proposal for evaluating educational research made by Michael Scriven in an informal discussion at a conference of the authors who collaborated on this yearbook. Scriven argued that in addition to the sections which customarily appear in educational research articles (e.g., "Introduction: Statement of Problem," "Methodology," "Results," "Discussion") there be added a section labeled

"Normative Implications." Here the author would analyze the morally significant effects of his research and the potential uses to which that research might be put. For example, could his findings be employed to train school youngsters to value or accommodate to "some win, most lose" competitive situations? Did his subjects learn, among other things, that knowledge should be sought because (or only if) it leads to prestige and privilege? The idea of requiring journal contributors to reflect openly about the normative implications of their work strikes me as a real liberation from the shallow dogma of "value-free" social science.[6] I would extend Scriven's proposal by suggesting that each study, each inquiry into what and why people learn, be expected to distinguish autonomous from nonautonomous (acts of) learning and to spell out whether its findings, or the potential applications thereof, enhance or restrict the scope of the former. In other words, we should ask ourselves: *Does the learning we are reporting and explaining take place at the cost of, or through extending, or with no effect on, the learners' capacities for autonomous activity?*[7]

Third, there is a related point which also concerns the evaluation of educational research. It is relatively well known that the question "What is he doing (did he do, etc.)?" is ambiguous. Does it inquire into "mere motion" (he is traveling southeast at three miles per hour) or into "intentional action" (he is going over to greet the Minister of Defense)? But this question is also ambiguous in a far less familiar way. Does it seek information on what our man is voluntarily or autonomously doing or only on what he is doing purposefully (but perhaps not voluntarily)? Neither of these two ambiguities in the way we talk and think should go unchallenged, especially when they appear in social science research. It is an implication of the science of autonomy that the unqualified use of terms like *act, behavior, learning, came to know, action, believes, desires, intends,* etc., should be viewed with as much suspicion as the unqualified use of such multimeaning terms as *law*, *objectivity*, *cause*, *model*, and *theory*. In brief, authors should be expected to indicate explicitly whether their observations, results, correlations, predictions, and so forth, apply to autonomous acts or to those that are purposeful but not autonomous or, again, to mere motions (or to say candidly that they do not know). Otherwise, what they are reporting and claiming will remain indeterminate and confusing.

---

[6] Arguments against this dogma are advanced in my introduction to Part VIII of *The Nature and Scope of Social Science*, pp. 691-703.

[7] An example of educational research which failed to raise this sort of question about its own methods and results is effectively criticized in Sarah Spinks' article, "Indians and Babies: Some OISE Projects," in *This Magazine Is About Schools*, vol. 4, no. 4 (Fall, 1970).

Last and most essential, the science of autonomy itself is an integral part of the domain of educational research. Autonomy can be learned by human beings, and it is, or ought to be, an aim of education. Thus, the new science is concerned with a distinctively human capacity which men learn and which education can foster; for this reason, it will obtain results which educational research cannot neglect. Moreover, the science of autonomy not only looks at *what* is learned, but inquires into the causal factors facilitating autonomous acts, i.e., by appealing to enabling conditions, REFs, antideterministic laws, and so forth. That is, the new science offers an alternative model for discovering *how and why* men learn what they do, at least in an important range of cases. From all of this it follows that educational research cannot hope to achieve one of its major goals—that of presenting an accurate and complete account of human learning—without developing a science of autonomy. In the absence of such a science, educational research can investigate only a restricted segment of what men learn and the kinds of causal factors which explain what they learn.[8]

---

[8] This paper has been revised from an earlier draft in order to take advantage of comments and criticisms made by the authors of the current volume and by my colleagues in philosophy at the University of Connecticut.

INDEX

Index

Ability grouping, assumptions underlying, 300
Accountability of schools for student behavior, rationale and problems of, 153-54
Achievement, analysis of, 198
Achievement tests: cultural and societal concepts in construction of, 194; factors in validity of, 198
Action: rational understanding of, 29-30; taking of, on judgment as stage in model building, 255; Weber's and Winch's understandings of, 44
Adult behavior, balance of interdependence and social responsiveness in, 159
Aesthetic domain: clarification of, 259; range of research activities relevant to, 258; relevance of structure of knowledge for research in, 273-74
Aesthetic experience: aspects of evaluation of, 263; tastes in, 259
Aesthetician, motive of, to move up explanatory ladder, 273
Aesthetic inquiry: clarification of strata of, 282; generic ways of failing in, 282; limits of, 258-87; nature of field of, 261
Aesthetic knowledge: derivation of, 264-66; ladder indicating structure of, 267
Aesthetic objects: disagreement about evaluations of, 266; interest of science and history in, 271
Aims of education, universally accepted types of, not measurable, 148-49
American Historical Association, 226
Analytic domains, differences between, and empirical domains, 134-36
Antideterministic uniformities, examples of, 352
Aristotle, 26, 79, 80, 86, 87, 88, 89, 195, 205, 312, 320
Arnheim, Rudolph, quoted, 275-76
Art: first job in evaluation of, 264; qualities of historical accounts of, 278
ASCD yearbook (1964), 105 ff
Assertions, relation of evidence to, 21
Assimilation, accommodation, and equilibration, views of processes of, 321-22
Assimilation, idea of, 308-9
Association for Supervision and Curriculum development, 165, 171, 177, 178, 183, 184, 186, 187, 189, 190, 191; quoted, 171
Attitudes: attempts at measurement of, in adults, 150-51; consideration of background in measurement of, 151; empirical test of behavior as evidence of, 152
Attitude tests, cultural and social concepts in construction of, 194
Autonomous actions, identification of, 331-38
Autonomous activity, effect of learning on, 354
Autonomous and nonautonomous acts, distinction between, 332-33
Autonomous beliefs, criteria for identification of, 334-35
Autonomous desires, criteria for identification of, 335-38
Autonomy, science of, 331-38
Avoidability, assumptions of, 328
Axiom of choice, 106

Bacon, Francis, 79, 195, 234
Bailyn, Bernard, 194
BDO (behaviorally defined objectives), 143, 144, 145, 146, 147, 152, 158, 160, 161
Beck, Robert H., 166
Behavioral objectives: dogma of, 219; nature of, 221
Behaviorally defined objective(s) (BDO): instances of confusion from use of terms of, 161; nature of, 143; practical concern of, 144-45; problems of attitude tests in re-

lation to, 152-53; relationship of researchers to schoolmen in relation to, 144; use of set of examples as, 166
Berensen, Bernard, 275
Berkeley, George, 196
Black, Max, 15; quoted, 15-16
Bridgman, P. W., 104, 197
Broudy, Harry S., 4
Buber, Martin, 333

Campbell, Donald T., 58, 59
Cantril, Hadley, 245
Causes, definition of, 128
Causes and reasons, 127-30
Central Midwestern Regional Educational Laboratory, inquiry of, 283
Closure, theorem of, 297
Cognition, basic considerations of, 106
Committee on the Economy of Time in Education (NEA), 228
Common sense, theory of, 57
Comparative education, approach to, 193
"Complete," a useful notion of, 124
Comprehension theorem, explanation of, 125
Computer simulation techniques, 64
Conant, James B., 141
Conceptual analysis, concern for, by philosophers, 258
Condorcet, Marquis de, 206
Controversial questions in the classroom, inquiries to be made about, 241-42
Council of the International Project for the Evaluation of Educational Achievement (IEA), 210, 211, 212, 213, 214, 215, 216; study of, 210 ff
Cremin, Lawrence A., 194
Criteria, cluster of, for research, 14-15
Criticism, effect of use of language in evaluation of object of, 266-67
Cross-national empirical studies, prior analysis needed for, 214
Cross-national studies, 210-16; value of, 193-94
Culturally deprived, use of term of, to evade responsibility, 174
Culture, implications of self-reproduction of, 290-91
Curriculum: attempt to establish on basis of replicative use of schooling, 228; incorporation of behaviors into, 221; objectivity and subjectivity in relation to, 140-41
Curriculum construction, assumptions underlying, 220-22
Curriculum development, problem of the inclusion of alternates in, 163

Darwin, Charles, 196, 291
Data-storage system, design of, 126-27
De Haan, Robert F., quoted with Doll, 172
Dennison, George, quoted, 349-50
Descartes, Rene, 206, 209
Developmental goals, setting of, by adults, 69
Deweyan model, neglect of, in education, 88
Dewey, John, 16, 82, 85, 86, 87, 88, 206, 208, 213, 242, 245, 311, 315, 317
Dewey's ends-in-view, nature of, 86
Dewey's theory of inquiry, aberrations arising from, 86-88
Diagnosis of situation, stage of, in model building, 247-49
Dietl, Paul, 178
Disadvantaged, redemption of school education for the, 300
Discourse: confusion arising from category mistakes in, 273; levels of, 267
Discovery, logic of, 134-38
Diverse learning pathways, implications of, 303-4
Doll, Ronald C., quoted with De Haan, 172
Dunkel, Harold B., 30

Early education, linkage of, with postnatal stages of embryological development, 328
Early environment, differential effect of, 301
Ecker, David W., 4
Eddington, Sir Arthur, 199
Education: improvement as aim of, 79; need of normative bases for, 93
Educational cycle and its closure, 289 ff
Educational experience, expansion of, 87
Educational journals, need for publication of ideas in, 141
Educational measures, dependence of, on definition of education, 198

Educational objectives: assumptions in behavioral definition of, 143-44; factors in choice of, 160-61
Educational objectives (non-BDO), rationale for use of, 153
Educational opportunity, cross-national studies of, 193
Educational policy, generation of, from notions of the nature of man, 26-27
Educational research: autonomy as part of domain of, 355; behaviorism in, 163; cluster of criteria for, 14-15; criteria for distinctiveness of, 23-24; coherence and self-consistency in, 199; current dogmas concerning, 258; possible directions for, 139; rationale for, found in clusters of theories, 205; reconstruction of, 23-24; relevance of, on educational psychology, 224; source of criteria for, 9-10; theoretical constructs in, 200-10
Educational studies, assertions as related to evidence in, 79
Educational theory, restraints upon, 289-90
Educators, military jargon of, 83
Effective instruction (IE), 168, 169, 171; conditions of, 168
Effectiveness, good intentions and accountability for, 144-48
Effectiveness of teaching, evaluation of, 158-59
Einstein, Albert, 199
Emphatic understanding, Weber on, 30-33
Empirical, definition of, 195
Empirical and experimental research, falseness of dichotomy between, 194-95
Empirical measurement, presupposition of theories in, 216
Empirical research, prior conceptual analysis of culture in, 216
Empiricist tradition, assumptions of, affirmed, 24-25
Enabling and sufficient conditions, relation of, to science of autonomy, 332
Enabling condition laws (EC), discussion of, 343 ff
Ends-in-view, Dewey's conception of, 86
Epistemology, rejection of philosophical justification of, 51
Equality: formulation of, 295; import of postulate of, 293-94; modifications of postulate of, 299; postulate of, in educational theory, 292-304; Rousseau's and Hobbes's models of, 293-95
Ethics, subject matter of, 131-32
Ethologists, global explanations by, 70
Euclid, 106
Evidence, need to specify form of, 146
Exemplificationality, definition of, 278
Expanding educational experience, Schwab's emphasis on, 87
Experience, description of, in associationist categories, 64-65
Experimental consequences: aspects of, of evaluative significance, 256-57
Experimental results, effect of affective tone of experiment, 59-60

Facts, acceptance of all of, by scientists, 47-49
Feyerabend, P. K., 118, 197
Fichte, Johann Gottlieb, 208
Formalism and intuitionism in cognition, 106
Formalist analysis, simple account of, 121-22
Formalists, need of, for intuitions, 107
Formal objectives, effect of, on student, 83-84
Frames of reference: construction of, as stage in model building, 249-52; exploration of questions in construction of, 252
Freedom: postulates of, 304-5; testing postulate of, 311; voluntary "negative" aspect of postulate of, 309
Freedom of the will, debates on, 304-5
Freud, Sigmund, 94

Gage, N. L., 4
General education: criteria for judging efficacy of, 236; inclusion of, in teacher education programs, 227; our schizophrenia about, 227
General theory, requirement of, 291
Gephart, William J., quoted with Bartos, 20
Gestaltist approach, use of, in personality assessment, 84
Gödel theorem, 106

Goodman, Nelson, 279, 280
Good intentions: discussion of, in instruction, 153-55; accountability in relation to, 148-53
Gowin, D. Bob, 3, 94, 103, 336
Guthrie, E. R., 59

Hanson, N. R., 48, 52
Hare, R. M., 91, 92
Harris, W. T., 82
Hart, H. L. A., 128
Hawkins, David, 4
Head Start, 175
Hegel, George Wilhelm Friedrich, 288, 295
Hempel, Carl, 86, 107, 122, 198
Hilgard, Ernest R., 49, 51
History: important learnings of, 115-16; nature of explanations in, 125
Hobbes, Thomas, 289, 290, 293, 294, 295, 296, 306, 307
Holmes, Brian, 4
Hull, Clark L., 62
Human agent, concept of, 314
Human behavior, understanding of, 27
Human freedom and science, alleged incompatibility of, 327-31
Human nature, scope of education in relation to, 287-326
Humanitarian concerns for well-being of student, 155
Human nature: characteristics of, implied by three generalizations, 288-89; examination of postulates of, for theory of education, 287 ff; postulates underlying, 287
Hume, David, 130, 196, 264, 319
Husen, Torsén, 212, 215
Hypothesis, controlling influence of, 197
Hypothetico-deductive methods, significant difference of induction to, 196-97
Hypothetico-deductive theory, discussion of, 196-97
Hypothetico-deductive theory of scientific method, implications of, 199-200

Ideologies, difference of, in ASCD and NSSE yearbooks, 172
Incompatibility dogma (ID): doctrine underlying, 328-29; nature of, 328; notion of boundary conditions of, 329
Individualization: different meanings of, 173; distinction between concepts of, 188 ff; review of treatment of, in NSSE and ASCD yearbooks, 165-72
Individualized instruction (II): two meanings of, 164-73 true meaning of, 177; views of, recapitulated, 183-84
Individualizing instruction, conclusions about, in NSSE yearbook, 165 ff
Individually Prescribed Instruction (IPI), values of, 176
Inferences and predictions, 119-27
Information-processing theories of learning, 64
Instruct: analysis of, 179-80; analysis of pedagogical sense of, 181-83
Instruction: concept of, 178; definition of, 182-83
International achievement tests, question of validity of, 212
International attitude tests, possible lack of validity of, 212
Interpretative use of schooling, centrality of, 225
Intersubjectivism: erroneous application of doctrine of, in research, 110; definitions in relation to, 104-5; fallacy of, 101; formatism and intuitionism in relation to, 106-12; evils of, 102-16
Introspectionism, result of reaction to, 98

James, William, 206
Jefferson, Thomas, 208
Job Corps, 175

Kaelin, Eugene F., 4
Kant, Immanuel, 314, 315, 316, 317, 319, 320, 323, 324, 325, 326
Kant's *Critiques*, analyis of, 316-17
Keller, Helen, 15
Kerlinger, F. N., 16, 17, 18, 78
Kessen, William, 69, 104
Keynes, John M., 221
Knowledge: implicit structure of, 266-67; ladder of structure of, 267
Krimerman, Leonard I., 4
Krupskaya, N. K., 207
Kuhn, T. S., 48, 50, 52, 118

Langer, Susanne, 274
Language, dangers in, as tool for judging art, 259-60
Language facility, role of, in reading and discussion, 237
Laplace, de, Pierre Simon, 112
Latent learning: explanation of, by drive-reduction theory of reinforcement, 67-68; phenomenon of, 65-68
Learning: agreement on evidence of 145; effect of space design on, 17; the jump from psychology of learning to, 31; ways used by person in, 221-22
Learning abilities, variance in, 298
Learning research, assumptions underlying, 220 ff
Learnings: controversy over question of, 58; three views of, 58-60; statements concerning three views of, 60-61
Learning theories, views of, 47-49
Lenin, Nikolai, 207, 208
Levitine, George, 270; quoted, 270
Life test-task, failure of school-learned items to function in, 237-38
Life uses of schooling, 219-38
Livingston, William, 245
Locke, John, 195, 206, 208, 209, 290
Lowenheim-Skolem theorem, 106

McCarthy, Mary, 260, 261, 272, 274, 276
McClellan, James E., 4
Makarenko, A. S., 207
Malcolm X, 15
Mandler, G., 104
Marx, Karl, 94, 100, 207, 208, 209, 295
Maslow, Abraham, 245, 333
May, Rollo, 333
Meaning, distinctions in interpretations of, 35-36
Means-end relationships: deliberation on, as stage in model building, 252-55; questions in deliberations of, 254-55
Medawar, P. B., 196
Mendel, Gregor Johann, 32, 296
Mendel's law, 32
Mental state, understanding of, 32-33
Metacritic, moves of, 277
Metacriticism: criteria for success in, 282-83; nature of, 267, 273; relevance of, to educational research, 280-84; role of, 266-67
Metatheory: nature of, 267-68, 273; recent approach to, 278
Mill, John Stuart, 195, 196, 197, 208
Milton, John, 236
Model: Dewey's theory of inquiry as example of, 85-86; means of judging of, 84-85; utility of, 85
Model for making and testing value judgments, 242 ff
Moral education, problems in developing materials for, 90-91
Moral imperatives, Hare's call for set of, 91-92
Morgan, Douglas N., 271; quoted, 271
Morgenstern, Oskar, 294
Morris, Charles, 248
Myrdal, Gunnar, 201

Narrow range of variation (NRV), 168, 169; principle of, 168
Neutral observation language, discussion of, 48-49
NSSE yearbook (1962), 164-65 ff
National Society for the Study of Education, 1, 3, 4, 164, 165, 168, 171, 177, 178, 183, 184, 186, 188, 191, 192
Neurath, Otto, 199
Newton, Isaac, 289
Nonautonomous action, identification of, 331-38
Nonbehavioral educational objectives, rationale for use of, 153
Normative base for education, need of education for, 93
Normative patterns, discussion of, 200-1
Normative research, appeal of, to educators, 90

Objectives, failure of behavior as measurement of attainment of, 149
Objectives in behavioral terms, problems of, 80
Objectivity: confusion in meanings of, 94-98; failure of neutral observation as guarantee of, 52; notion of intersubjectivism as criterion of, 99
Objectivity-subjectivity: background of confusion relating to, 98-100; confusion in contrasts of, 95-96; foundations of, 94-116; methodological applications in relation to, 116-30

Observable behavior, requirement of, for testing, 161
Observables, inability to find unique set of, 55
Observable terms, role of, in operational definitions, 54
Observational categories, meaning of, 49-50
Observation language, distinction between, and reports of facts, 47
Occam, 99
Office of Economic Opportunity (OEO), 175, 176, 177; programmed instruction not features in programs of, 173-74
Office of Education (USOE), 5, 141
Office of Strategic Services (OSS), 84
Operational definitions: nature of, 54; relevance of, to theory-dependency thesis, 53-56; roles of, 54-55
Operationalism: commitment of, to nonoperationalizable definitions, 100; discussion of, 197-98

Pareto, V., 200
Pavlov, Ivan P., 99, 207
Peirce, C. S., 206, 267
Pepper, Stephen C., 276, 281; work of, 281
Perception, problems of, 52-53
Petrie, Hugh G., 3, 48, 52
Phenomenological metacriticism, theoretical justification of, 277
Phenomenology, role of, 274-75
Philosophers, concern of, for value judgments and process of their development, 240
Philosophical modes of inquiry, criticisms of, 194
Philosophy, lack of interest in "nonscientific" parts of, 78
Philosophy and educational research, convergence of, in area of science, 77
Philosophy of science, debates in, 195-200
Piaget, Jean, 290, 307, 308, 313, 319, 321, 323, 326
Piaget's conceptual trinity, 313
Picasso, *Guernica* of, 264; *Night Fishing at Antibes* of, 268
Plato, 205, 206, 208, 312
Polanyi, Michael, 223, 229, 231; quoted, 230
Popper, K. R., 120, 121, 134, 136, 196, 198, 200
Practice, effect of, on later performance, 56
Pragmatic view of values, common objections to, 132
Prall, D. W., 268
Prediction and explanation, logical differences in the enterprise of, 123
Predictions and inferences, 119-27
Pretest, essential nature of, 56
Problem-structure of development, controversies in relation to, 70
Programmed instruction (PI), 175-176, 177; compelling argument for, in War on Poverty, 174-75; factor of, in individualization, 164; impact of, on research activity, 174; role of, 173; values of, 176
Public schools, nature of conflict of, with OEO, 176-77

Quesnay, François, 291
Quine, W. V. O., 48, 50, 52

Range-extending factor (REF), discussion of, 349-50
Rationality, 312-26; postulate of, 312
Rational understanding, explanation of notion of, 34
Readiness, inequalities explained in differences in, 298
Reasoning, research in strategies of, 129-30
Reasons-causes nexus, ramifications of, in social sciences, 128
Reid, L. A., 210
Reinforcement: drive-reduction view of, 66; factors in, 68-69; type of, 82
Replicative use of schooling, doctrine of, 225
Reporting an experience, effect of factors on, 262-63
Research: examples of, 11-14; narrow view of, 78-79; programmed instruction and OEO programs in relation to, 173-74; restriction of, by dogma of behavioral objectives, 219
Research and tacit knowing, 232-38
Research generalizations, effect of changing educational patterns on, 10
Research Institute of Comparative Education and Culture (Kyushu),

210, 211, 212, 214; study of moral education by, 210 ff
Research methodology, example of, 56
Research practice, methodological strictures in, 72
"Residues," motivation of behavior by, 202
Rogers, Carl, 245, 328, 333
Role-playing; problem of, in attitude responses, 151-52; student's use of, in behavioral tests, 150
Rosenberg, Harold, 260, 261, 272, 274, 276; quoted, 260, 272
Rosenthal, R., 57
Roszak, T., 46
Rousseau, Jean J., 288, 295
Russell, Bertrand, 112

Santayana, George, 268; quoted 269-70
Sartre, Jean Paul, 207, 333
Scheffler, Israel, 197, 199
Schooling: argument for interpretive use of, 225; indications of functionality of, 237; judgment of, in terms of, aims, 148; life uses of, 220-25; research on life uses of, 219; testing assumptions underlying, 220-21; uses of, 220-23
School input, nature of, 224
School learnings: doctrines as to nature of, 222; requirements for testing functionality of, 237-38; research and life uses of, 219-20; theory that use of, in life is by replication, 228-29
Schools, success of, evaluated without regard to student behavior, 153-54
Schwab, Joseph J., 16, 87
Science: characteristics of, 80; impact of, on concept of nature of man, 26-27; definition of, as value-free, 79; limited utility of, for education, 79; nature of, 138-39; view of, as partly descriptive, 79
Science and human freedom, alleged incompatibility of, 327-31
Science of autonomy; explanation of, 338 ff; implications of, for educational research, 352-53; major distinctions underlying, 332-33
Scientific methods, implications of hypothetico-deductive method for, 199-200
Scientific theory, purpose of, 296
Scientists, challenge to acceptance of all facts by, 47-49
Scriven, Michael, 3, 355
Sense-data theories, problems of, 53
Shakespeare, William, 206
Shulman, Lee S., 23; quoted, 21, 22
Sign-learning theory (Tolman), 59
Skinner, B. F., 46, 60, 66, 69, 101, 175, 327; quoted, 327
Skinner's operant analysis, 60
Smith, Adam, 291
Snow, Richard E., 139
Social-class background, cross-national studies of, 193
Social research, reasons-causes nexus in, 128-30
Social science: relevance of approaches to, 240; *values* as subject matter of, 239-40
Space design, effect of, on learning, 17
Specific-response laws, reliance of scientific inquiry on, 330
Spinoza, Baruch, 306, 307
Stanford University, symposium at, 23-24
Stated objectives: difficulties involved in use of, 81-82; irrelevance of, 81; relevant characteristics, 81
Statistical tendency statements, understanding of, as approximations of universal laws, 330-31
Steinberg, Ira S., 4
Stevens, S. S., 19
Strike, Kenneth A., 3
Strong knowledge claim, nature of, 113-19
Students rights, limits of, 156
Subjective probability, discussion of, 110-12
Subject matter: nature of, 267; nature of types of, 23
Sugarman, B., 92
Swenson, Esther J., quoted, 167
Symbolic systems, properties of, 280

Tacit knowing, hypothesis of, 232
Tacit knowing and research, 232-38
Teach, standard paradigm of, 178-79
Teacher education, convictions as to content of, 226-27
Teacher effectiveness, student performance as measure of, 146
Teaching: conceptions of, analyzed, 2; factors other than effectiveness in

judgment of, 162-63; nature and purpose of, 157; presumption of good intentions in, 157-58
Teaching and learning, aims of empirical research in relation to, 156-57
Teaching effectiveness, realization of intentions necessary for, 147
Testing, lack of validity in, 226-27
Testing value judgments, 139-57
Tests (international), problems involved in construction of, 212-14
Theoretical constructs: criteria of coherence and comprehensiveness in, 205-6; establishment of, 200-4; sources for derivation of, 208-9; values of, 209
Theoretical perspectives, effect of, on observation, 48
Theories, testing of, by observing facts, 47 ff
Theories of learning, lack of relationship of, to learning, 72
Theories vs. observations, 117-19
Theory, importance of, in underdeveloped fields, 288; Whorfian version of, 50
Theory-dependency thesis of observations: extreme form of, 50; relevance of operational definitions to, 53-56
Theory-dependency of observation thesis, illustrations of, in five areas, 58 ff
Theory-dependency thesis, extreme form of, 50-51
Thomas, Lawrence G., 4
Thorndike, E. L., 59; quoted, 175, 192, 220, 221
Tolman, E. C., 59
Toulmin, Stephan, 48, 52, 112
Transformational linguists, meaning of competence to, 65
Travers, Robert M. W., 10, 18

Undergoing conflict, stage of, in making judgments, 244-47
Understanding: crucial feature of all tests of, 126; explanations as vehicles of, 125
Understanding (an action), concepts of Winch and Weber relating to, 40-43
Understanding as tacit, 228-32
Universal rules, appeal to self-evident principles for claim of knowledge of, 262-63
Upward Bound, 175
Usefulness, criterion of, 202
U.S. Office of Education, 141

Value, pragmatic view of, 132
Value claims: objective and subjective nature of, 131; problem of verification of, 96-97
Value judgments: behavioral clue of nature of, 257; empirical studies relating to, 239; enhancement of quality of, in education, 242; proposed model for making and testing of, 242-57; scientific inquiry into areas of concern to makers of, 240-41; stages of testing and making of, 243-44; testing of, 139-57; undergoing conflict as stimulus for making of, 244-47
Values, communication between philosophers and social scientists on, 239; pragmatic view of, 132
Verification, meaning of independent means of, 109
Verstehen, 27, 28, 31, 38, 40, 127; disputes clustering about notion of, 28; doctrine of, 37; notion of, in writings of Weber, 27-28; quality of, 61; radical empiricist views on, 37-38; Weber's discussion of, 28-29
Vienna Circle, 199
von Neumann, John, 294

Weak knowledge-claim, nature of, 114
Weber, Max, 28, 29, 30, 31, 32, 33, 34, 35, 38, 39, 42, 43, 44, 45, 61, 77; quoted, 31, 35, 36, 37; criticism of, 34-35; failure of critics of, to note his distinctions, 40; views of critics of, 38-39
Weitz, Morris, 278, 279, 280
Wertheimer, Max, 316
Whitehead, A. N., 84
Williams, N., 92
Wilson, J., 90
Winch, Peter, 39, 40, 44, 45; quoted, 41-42, 43; concept of understanding of, 43-45; objection to analysis of, 43
Wittgenstein, Ludwig, 63, 276

Ziff, Paul, 274

INFORMATION CONCERNING THE NATIONAL SOCIETY FOR THE STUDY OF EDUCATION

1. Purpose. The purpose of the National Society is to promote the investigation and discussion of educational questions. To this end it holds an annual meeting and publishes a series of yearbooks.

2. Eligibility to Membership. Any person who is interested in receiving its publications may become a member by sending to the Secretary-Treasurer information concerning name, title, and address, and a check for $9.00 (see Item 5), except that graduate students, on the recommendation of a faculty member, may become members by paying $7.00 for the first year of their membership. Dues for all subsequent years are the same as for other members (see Item 4).

Membership is not transferable; it is limited to individuals, and may not be held by libraries, schools, or other institutions, either directly or indirectly.

3. Period of Membership. Applicants for membership may not date their entrance back of the current calendar year, and all memberships terminate automatically on December 31, unless the dues for the ensuing year are paid as indicated in Item 6.

4. Duties and Privileges of Members. Members pay dues of $8.00 annually, receive a cloth-bound copy of each publication, are entitled to vote, to participate in discussion, and (under certain conditions) to hold office. The names of members are printed in the yearbooks.

Persons who are sixty years of age or above may become life members on payment of fee based on average life-expectancy of their age group. For information, apply to Secretary-Treasurer.

5. Entrance Fee. New members are required the first year to pay, in addition to the dues, an entrance fee of one dollar.

6. Payment of Dues. Statements of dues are rendered in October for the following calendar year. Any member so notified whose dues remain unpaid on January 1 thereby loses his membership and can be reinstated only by paying a reinstatement fee of fifty cents.

School warrants and vouchers from institutions must be accompanied by definite information concerning the name and address of the person for whom membership fee is being paid. Statements of dues are rendered on our own form only. The Secretary's office cannot undertake to fill out special invoice forms of any sort or to affix notary's affidavit to statements or receipts.

Cancelled checks serve as receipts. Members desiring an additional receipt must enclose a stamped and addressed envelope therefor.

7. Distribution of Yearbooks to Members. The yearbooks, ready prior to each February meeting, will be mailed from the office of the distributor, only to members whose dues for that year have been paid. Members who desire yearbooks prior to the current year must purchase them directly from the distributor (see Item 8).

8. Commercial Sales. The distribution of all yearbooks prior to the current year, and also of those of the current year not regularly mailed to members in exchange for their dues, is in the hands of the distributor, not of the Secretary. For such commercial sales, communicate directly with the University of Chicago Press, Chicago, Illinois 60637, which will gladly send a price list covering all the publications of this Society. This list is also printed in the yearbook.

9. Yearbooks. The yearbooks are issued about one month before the February meeting. They comprise from 600 to 800 pages annually. Unusual effort has been made to make them, on the one hand, of immediate practical value, and, on the other hand, representative of sound scholarship and scientific investigation.

10. Meetings. The annual meeting, at which the yearbooks are discussed, is held, as a rule, in February at the same time and place as the meeting of the American Association of School Administrators. Members will be notified of other meetings.

Applications for membership will be handled promptly at any time on receipt of name and address, together with check for $9.00 (or $8.50 for reinstatement). Applications entitle the new members to the yearbook slated for discussion during the calendar year the application is made.

Herman G. Richey, *Secretary-Treasurer*

5835 Kimbark Ave.
Chicago, Illinois 60637

PUBLICATIONS OF THE NATIONAL SOCIETY FOR THE STUDY OF EDUCATION

NOTICE: Many of the early Yearbooks of this series are now out of print. In the following list, those titles to which an asterisk is prefixed are not available for purchase.

POSTPAID PRICE

*First Yearbook, 1902, Part I—*Some Principles in the Teaching of History.* Lucy M. Salmon

*First Yearbook, 1902, Part II—*The Progress of Geography in the Schools.* W. M. Davis and H. M. Wilson

*Second Yearbook, 1903, Part I—*The Course of Study in History in the Common School.* Isabel Lawrence, C. A. McMurry, Frank McMurry, E. C. Page, and E. J. Rice

*Second Yearbook, 1903, Part II—*The Relation of Theory to Practice in Education.* M. J. Holmes, J. A. Keith, and Levi Seeley

*Third Yearbook, 1904, Part I—*The Relation of Theory to Practice in the Education of Teachers.* John Dewey, Sarah C. Brooks, F. M. McMurry, *et al.*

*Third Yearbook, 1904, Part II—*Nature Study.* W. S. Jackman

*Fourth Yearbook, 1905, Part I—*The Education and Training of Secondary Teachers.* E. C. Elliott, E. G. Dexter, M. J. Holmes, *et al.*

*Fourth Yearbook, 1905, Part II—*The Place of Vocational Subjects in the High-School Curriculum.* J. S. Brown, G. B. Morrison, and Ellen Richards

*Fifth Yearbook, 1906, Part I—*On the Teaching of English in Elementary and High Schools.* G. P. Brown and Emerson Davis

*Fifth Yearbook, 1906, Part II—*The Certification of Teachers.* E. P. Cubberley

*Sixth Yearbook, 1907, Part I—*Vocational Studies for College Entrance.* C. A. Herrick, H. W. Holmes, T. deLaguna, V. Prettyman, and W. J. S. Bryan

*Sixth Yearbook, 1907, Part II—*The Kindergarten and Its Relation to Elementary Education.* Ada Van Stone Harris, E. A. Kirkpatrick, Marie Kraus-Boelté, Patty S. Hill, Harriette M. Mills, and Nina Vandewalker

*Seventh Yearbook, 1908, Part I—*The Relation of Superintendents and Principals to the Training and Professional Improvement of Their Teachers.* Charles D. Lowry

*Seventh Yearbook, 1908, Part II—*The Co-ordination of the Kindergarten and the Elementary School.* B. J. Gregory, Jennie B. Merrill, Bertha Payne, and Margaret Giddings

*Eighth Yearbook, 1909, Part I—*Education with Reference to Sex: Pathological, Economic, and Social Aspects.* C. R. Henderson

*Eighth Yearbook, 1909, Part II—*Education with Reference to Sex: Agencies and Methods.* C. R. Henderson and Helen C. Putnam

*Ninth Yearbook, 1910, Part I—*Health and Education.* T. D. Wood

*Ninth Yearbook, 1910, Part II—*The Nurses in Education.* T. D. Wood, *et al.*

*Tenth Yearbook, 1911, Part I—*The City School as a Community Center.* H. C. Leipziger, Sarah E. Hyre, R. D. Warden, C. Ward Crampton, E. W. Stitt, E. J. Ward, Mrs. E. C. Grice, and C. A. Perry

*Tenth Yearbook, 1911, Part II—*The Rural School as a Community Center.* B. H. Crocheron, Jessie Field, F. W. Howe, E. C. Bishop, A. B. Graham, O. J. Kern, M. T. Scudder, and B. M. Davis

*Eleventh Yearbook, 1912, Part I—*Industrial Education: Typical Experiments Described and Interpreted.* J. F. Barker, M. Bloomfield, B. W. Johnson, P. Johnson, L. M. Leavitt, G. A. Mirick, M. W. Murray, C. F. Perry, A. L. Stafford, and H. B. Wilson

*Eleventh Yearbook, 1912, Part II—*Agricultural Education in Secondary Schools.* A. C. Monahan, R. W. Stimson, D. J. Crosby, W. H. French, H. F. Button, F. R. Crane, W. R. Hart, and G. F. Warren

*Twelfth Yearbook, 1913, Part I—*The Supervision of City Schools.* Franklin Bobbitt, J. W. Hall, and J. D. Wolcott

*Twelfth Yearbook, 1913, Part II—*The Supervision of Rural Schools.* A. C. Monahan, L. J. Hanifan, J. E. Warren, Wallace Lund, U. J. Hoffman, A. S. Cook, E. M. Rapp, Jackson Davis, and J. D. Wolcott

*Thirteenth Yearbook, 1914, Part I—*Some Aspects of High-School Instruction and Administration.* H. C. Morrison, E. R. Breslich, W. A. Jessup, and L. D. Coffman

*Thirteenth Yearbook, 1914, Part II—*Plans for Organizing School Surveys, with a Summary of Typical School Surveys.* Charles H. Judd and Henry L. Smith

*Fourteenth Yearbook, 1915, Part I—*Minimum Essentials in Elementary School Subjects—Standards and Current Practices.* H. B. Wilson, H. W. Holmes, F. E. Thompson, R. G. Jones, S. A. Courtis, W. S. Gray, F. N. Freeman, H. C. Pryor, J. F. Hosic, W. A. Jessup, and W. C. Bagley

*Fourteenth Yearbook, 1915, Part II—*Methods for Measuring Teachers' Efficiency.* Arthur C. Boyce

*Fifteenth Yearbook, 1916, Part I—*Standards and Tests for the Measurement of the Efficiency of Schools and School Systems.* G. D. Strayer, Bird T. Baldwin, B. R. Buckingham, F. W. Ballou, D. C. Bliss, H. G. Childs, S. A. Courtis, E. P. Cubberley, C. H. Judd, George Melcher, E. E. Oberholtzer, J. B. Sears, Daniel Starch, M. R. Trabue, and G. M. Whipple

*Fifteenth Yearbook, 1916, Part II—*The Relationship between Persistence in School and Home Conditions.* Charles E. Holley

POSTPAID PRICE

*Fifteenth Yearbook, 1916, Part III—*The Junior High School.* Aubrey A. Douglas..

*Sixteenth Yearbook, 1917, Part I—*Second Report of the Committee on Minimum Essentials in Elementary-School Subjects.* W. C. Bagley, W. W. Charters, F. N. Freeman, W. S. Gray, Ernest Horn, J. H. Hoskinson, W. S. Monroe, C. F. Munson, H. C. Pryor, L. W. Rapeer, G. M. Wilson, and H. B. Wilson...........

*Sixteenth Yearbook, 1917, Part II—*The Efficiency of College Students as Conditioned by Age at Entrance and Size of High School.* B. F. Pittenger................

*Seventeenth Yearbook, 1918, Part I—*Third Report of the Committee on Economy of Time in Education.* W. C. Bagley, B. B. Bassett, M. E. Branom, Alice Camerer, J. E. Dealey, C. A. Ellwood, E. B. Greene, A. B. Hart, J. F. Hosic, E. T. Housh, W. H. Mace, L. R. Marston, H. C. McKown, H. E. Mitchell, W. C. Reavis, D. Snedden, and H. B. Wilson..

*Seventeenth Yearbook, 1918, Part II—*The Measurement of Educational Products.* E. J. Ashbaugh, W. A. Averill, L. P. Ayers, F. W. Ballou, Edna Bryner, B. R. Buckingham, S. A. Courtis, M. E. Haggerty, C. H. Judd, George Melcher, W. S. Monroe, E. A. Nifenecker, and E. L. Thorndike.........................

*Eighteenth Yearbook, 1919, Part I—*The Professional Preparation of High-School Teachers.* G. N. Cade, S. S. Colvin, Charles Fordyce, H. H. Foster, T. S. Gosling, W. S. Gray, L. V. Koos, A. R. Mead, H. L. Miller, F. C. Whitcomb, and Clifford Woody . . . .

*Eighteenth Yearbook, 1919, Part II—*Fourth Report of Committee on Economy of Time in Education.* F. C. Ayer, F. N. Freeman, W. S. Gray, Ernest Horn, W. S. Monroe, and C. E. Seashore..

*Nineteenth Yearbook, 1920, Part I—*New Materials of Instruction.* Prepared by the Society's Committee on Materials of Instruction.........................

*Nineteenth Yearbook, 1920, Part II—*Classroom Problems in the Education of Gifted Children.* T. S. Henry...

*Twentieth Yearbook, 1921, Part I—*New Materials of Instruction.* Second Report by Society's Committee . . ..

*Twentieth Yearbook, 1921, Part II—*Report of the Society's Committee on Silent Reading.* M. A. Burgess, S. A. Courtis, C. E. Germane, W. S. Gray, H. A. Greene, Regina R. Heller, J. H. Hoover, J. A. O'Brien, J. L. Packer, Daniel Starch, W. W. Theisen, G. A. Yoakam, and representatives of other school systems...........

*Twenty-first Yearbook, 1922, Parts I and II—*Intelligence Tests and Their Use.* Part I—*The Nature, History, and General Principles of Intelligence Testing.* E. L. Thorndike, S. S. Colvin, Harold Rugg, G. M. Whipple. Part II—*The Administrative Use of Intelligence Tests.* H. W. Holmes, W. K. Layton, Helen Davis, Agnes L. Rogers, Rudolf Pintner, M. R. Trabue, W. S. Miller, Bessie L. Gambrill, and others. The two parts are bound together.................................

*Twenty-second Yearbook, 1923, Part I—*English Composition: Its Aims, Methods and Measurements.* Earl Hudelson...

*Twenty-second Yearbook, 1923, Part II—*The Social Studies in the Elementary and Secondary School.* A. S. Barr, J. J. Coss, Henry Harap, R. W. Hatch, H. C. Hill, Ernest Horn, C. H. Judd, L. C. Marshall, F. M. McMurry, Earle Rugg, H. O. Rugg, Emma Schweppe, Mabel Snedaker, and C. W. Washburne.............

*Twenty-third Yearbook, 1924, Part I—*The Education of Gifted Children.* Report of the Society's Committee. Guy M. Whipple, Chairman......................

*Twenty-third Yearbook, 1924, Part II—*Vocational Guidance and Vocational Education for Industries.* A. H. Edgerton and Others..........................

*Twenty-fourth Yearbook, 1925, Part I—*Report of the National Committee on Reading.* W. S. Gray, Chairman, F. W. Ballou, Rose L. Hardy, Ernest Horn, Francis Jenkins, S. A. Leonard, Estaline Wilson, and Laura Zirbes....................

*Twenty-fourth Yearbook, 1925, Part II—*Adapting the Schools to Individual Differences.* Report of the Society's Committee. Carleton W. Washburn, Chairman....

*Twenty-fifth Yearbook, 1926, Part I—*The Present Status of Safety Education.* Report of the Society's Committee. Guy M. Whipple, Chairman.....................

*Twenty-fifth Yearbook, 1926, Part II—*Extra-Curricular Activities.* Report of the Society's Committee. Leonard V. Koos, Chairman...........................

*Twenty-sixth Yearbook, 1927, Part I—*Curriculum-making: Past and Present.* Report of the Society's Committee. Harold O. Rugg, Chairman....................

*Twenty-sixth Yearbook, 1927, Part II—*The Foundations of Curriculum-making.* Prepared by individual members of the Society's Committee. Harold O. Rugg, Chairman

*Twenty-seventh Yearbook, 1928, Part I—*Nature and Nurture: Their Influence upon Intelligence.* Prepared by the Society's Committee. Lewis M. Terman, Chairman..

*Twenty-seventh Yearbook, 1928, Part II—*Nature and Nurture: Their Influence upon Achievement.* Prepared by the Society's Committee. Lewis M. Terman, Chairman.

Twenty-eighth Yearbook, 1929, Parts I and II—*Preschool and Parental Education.* Part I—*Organization and Development.* Part II—*Research and Method.* Prepared by the Society's Committee. Lois H. Meek, Chairman. Bound in one volume. Cloth $5.00

	POSTPAID PRICE
*Twenty-ninth Yearbook, 1930, Parts I and II—*Report of the Society's Committee on Arithmetic*. Part I—*Some Aspects of Modern Thought on Arithmetic*. Part II—*Research in Arithmetic*. Prepared by the Society's Committee. F. B. Knight, Chairman. Bound in one volume. Cloth	
Thirtieth Yearbook, 1931, Part I—*The Status of Rural Education*. First Report of the Society's Committee on Rural Education. Orville G. Brim, Chairman. Cloth	$2.50
Thirtieth Yearbook, 1931, Part II—*The Textbook in American Education*. Report of the Society's Committee on the Textbook. J. B. Edmonson, Chairman. Cloth	2.50
Paper	1.75
*Thirty-first Yearbook, 1932, Part I—*A Program for Teaching Science*. Prepared by the Society's Committee on the Teaching of Science. S. Ralph Powers, Chairman.	
Paper	
Thirty-first Yearbook, 1932, Part II—*Changes and Experiments in Liberal-Arts Education*. Prepared by Kathryn McHale, with numerous collaborators. Cloth	3.00
*Thirty-second Yearbook, 1933—*The Teaching of Geography*. Prepared by the Society's Committee on the Teaching of Geography. A. E. Parkins, Chairman. Cloth	
Paper	3.75
Thirty-third Yearbook, 1934, Part I—*The Planning and Construction of School Buildings*. Prepared by the Society's Committee on School Buildings. N. L. Engelhardt, Chairman. Cloth	2.50
Thirty-third Yearbook, 1934, Part II—*The Activity Movement*. Prepared by the Society's Committee on the Activity Movement. Lois Coffey Mossman, Chairman. Cloth	3.00
Paper	2.25
Thirty-fourth Yearbook, 1935—*Educational Diagnosis*. Prepared by the Society's Committee on Educational Diagnosis. L. J. Brueckner, Chairman. Cloth	5.00
Paper	3.75
*Thirty-fifth Yearbook, 1936, Part I—*The Grouping of Pupils*. Prepared by the Society's Committee. W. W. Coxe, Chairman. Cloth	
*Thirty-fifth Yearbook, 1936, Part II—*Music Education*. Prepared by the Society's Committee. W. L. Uhl, Chairman. Cloth	
*Thirty-sixth Yearbook, 1937, Part I—*The Teaching of Reading*. Prepared by the Society's Committee. W. S. Gray, Chairman. Cloth	
*Thirty-sixth Yearbook, 1937, Part II—*International Understanding through the Public-School Curriculum*. Prepared by the Society's Committee. I. L. Kandel, Chairman. Paper	
*Thirty-seventh Yearbook, 1938, Part I—*Guidance in Educational Institutions*. Prepared by the Society's Committee. G. N. Kefauver, Chairman. Cloth	
*Thirty-seventh Yearbook, 1938, Part II—*The Scientific Movement in Education*. Prepared by the Society's Committee. F. N. Freeman, Chairman. Paper	
*Thirty-eighth Yearbook, 1939, Part I—*Child Development and the Curriculum*. Prepared by the Society's Committee. Carleton Washburne, Chairman. Paper	
Thirty-eighth Yearbook, 1939 Part II—*General Education in the American College*. Prepared by the Society's Committee. Alvin Eurich, Chairman. Cloth	2.75
*Thirty-ninth Yearbook, 1940, Part I—*Intelligence: Its Nature and Nurture. Comparative and Critical Exposition*. Prepared by the Society's Committee. G. D. Stoddard, Chairman. Cloth	
*Thirty-ninth Yearbook, 1940, Part II—*Intelligence: Its Nature and Nurture. Original Studies and Experiments*. Prepared by the Society's Committee. G. D. Stoddard, Chairman. Cloth	
*Fortieth Yearbook, 1941—*Art in American Life and Education*. Prepared by the Society's Committee. Thomas Munro, Chairman. Paper	
Forty-first Yearbook, 1942, Part I—*Philosophies of Education*. Prepared by the Society's Committee. John S. Brubacher, Chairman. Cloth	4.00
Paper	3.25
Forty-first Yearbook, 1942, Part II—*The Psychology of Learning*. Prepared by the Society's Committee. T. R. McConnell, Chairman. Cloth	4.50
Paper	3.75
Forty-second Yearbook, 1943, Part I—*Vocational Education*. Prepared by the Society's Committee. F. J. Keller, Chairman. Cloth	3.25
Paper	2.50
Forty-second Yearbook, 1943, Part II—*The Library in General Education*. Prepared by the Society's Committee. L. R. Wilson, Chairman. Cloth	4.00
*Forty-third Yearbook, 1944, Part I—*Adolescence*. Prepared by the Society's Committee. Harold E. Jones, Chairman. Cloth	3.25
Paper	

	POSTPAID PRICE
Forty-third Yearbook, 1944, Part II—*Teaching Language in the Elementary School.* Prepared by the Society's Committee. M. R. Trabue, Chairman. Cloth	$3.00
*Forty-fourth Yearbook, 1945, Part I—*American Education in the Postwar Period: Curriculum Reconstruction.* Prepared by the Society's Committee. Ralph W. Tyler, Chairman. Paper	
*Forty-fourth Yearbook, 1945, Part II—*American Education in the Postwar Period: Structural Reorganization.* Prepared by the Society's Committee. Bess Goodykoontz, Chairman. Paper	
*Forty-fifth Yearbook, 1946, Part I—*The Measurement of Understanding.* Prepared by the Society's Committee. William A. Brownell, Chairman. Paper	
Forty-fifth Yearbook, 1946, Part II—*Changing Conceptions in Educational Administration.* Prepared by the Society's Committee. Alonzo G. Grace, Chairman. Cloth	4.50
*Forty-sixth Yearbook, 1947, Part I—*Science Education in American Schools.* Prepared by the Society's Committee. Victor H. Noll, Chairman. Paper	
Forty-sixth Yearbook, 1947, Part II—*Early Childhood Education.* Prepared by the Society's Committee. N. Searle Light, Chairman. Cloth Paper	4.50 3.75
Forty-seventh Yearbook, 1948, Part I—*Juvenile Delinquency and the Schools.* Prepared by the Society's Committee. Ruth Strang, Chairman. Cloth	4.50
Forty-seventh Yearbook, 1948, Part II—*Reading in the High School and College.* Prepared by the Society's Committee. William S. Gray, Chairman. Cloth Paper	4.50 3.75
Forty-eighth Yearbook, 1949, Part I—*Audio-visual Materials of Instruction.* Prepared by the Society's Committee. Stephen M. Corey, Chairman. Cloth	4.50
Forty-eighth Yearbook, 1949, Part II—*Reading in the Elementary School.* Prepared by the Society's Committee. Arthur I. Gates, Chairman. Cloth Paper	4.50 3.75
*Forty-ninth Yearbook, 1950, Part I—*Learning and Instruction.* Prepared by the Society's Committee. G. Lester Anderson, Chairman. Cloth Paper	 3.75
*Forty-ninth Yearbook, 1950, Part II—*The Education of Exceptional Children.* Prepared by the Society's Committee. Samuel A. Kirk, Chairman. Cloth Paper	 3.75
Fiftieth Yearbook, 1951, Part I—*Graduate Study in Education.* Prepared by the Society's Board of Directors. Ralph W. Tyler, Chairman. Cloth Paper	4.50 3.75
Fiftieth Yearbook, 1951, Part II—*The Teaching of Arithmetic.* Prepared by the Society's Committee. G. T. Buswell, Chairman. Cloth Paper	4.50 3.75
Fifty-first Yearbook, 1952, Part I—*General Education.* Prepared by the Society's Committee: T. R. McConnell, Chairman. Cloth Paper	4.50 3.75
Fifty-first Yearbook, 1952, Part II—*Education in Rural Communities.* Prepared by the Society's Committee. Ruth Strang, Chairman. Cloth Paper	4.50 3.75
Fifty-second Yearbook, 1953, Part I—*Adapting the Secondary-School Program to the Needs of Youth.* Prepared by the Society's Committee: William G. Brink, Chairman. Cloth Paper	4.50 3.75
Fifty-second Yearbook, 1953, Part II—*The Community School.* Prepared by the Society's Committee. Maurice F. Seay, Chairman. Cloth Paper	4.50 3.75
Fifty-third Yearbook, 1954, Part I—*Citizen Co-operation for Better Public Schools.* Prepared by the Society's Committee. Edgar L. Morphet, Chairman. Cloth Paper	4.50 3.75
Fifty-third Yearbook, 1954, Part II—*Mass Media and Education.* Prepared by the Society's Committee. Edgar Dale, Chairman. Cloth Paper	4.50 3.75
Fifty-fourth Yearbook, 1955, Part I—*Modern Philosophies and Education.* Prepared by the Society's Committee. John S. Brubacher, Chairman. Cloth Paper	4.50 3.75
Fifty-fourth Yearbook, 1955, Part II—*Mental Health in Modern Education.* Prepared by the Society's Committee. Paul A. Witty, Chairman. Cloth Paper	4.50 3.75
*Fifty-fifth Yearbook, 1956, Part I—*The Public Junior College.* Prepared by the Society's Committee. B. Lamar Johnson, Chairman. Cloth	

	POSTPAID PRICE
Fifty-fifth Yearbook, 1956, Part II—*Adult Reading*. Prepared by the Society's Committee. David H. Clift, Chairman. Cloth	$4.50
Paper	3.75
Fifty-sixth Yearbook, 1957, Part I—*In-service Education of Teachers, Supervisors, and Administrators*. Prepared by the Society's Committee. Stephen M. Corey, Chairman. Cloth	4.50
Paper	3.75
Fifty-sixth Yearbook, 1957, Part II—*Social Studies in the Elementary School*. Prepared by the Society's Committee. Ralph C. Preston, Chairman. Cloth	4.50
Paper	3.75
Fifty-seventh Yearbook, 1958, Part I—*Basic Concepts in Music Education*. Prepared by the Society's Committee. Thurber H. Madison, Chairman. Cloth	4.50
Fifty-seventh Yearbook, 1958, Part II—*Education for the Gifted*. Prepared by the Society's Committee. Robert J. Havighurst, Chairman. Cloth	4.50
Paper	3.75
Fifty-seventh Yearbook, 1958, Part III—*The Integration of Educational Experiences*. Prepared by the Society's Committee. Paul L. Dressel, Chairman. Cloth	4.50
Fifty-eighth Yearbook, 1959, Part I—*Community Education: Principles and Practices from World-wide Experience*. Prepared by the Society's Committee. C. O. Arndt, Chairman. Cloth	4.50
Paper	3.75
Fifty-eighth Yearbook, 1959, Part II—*Personnel Services in Education*. Prepared by the Society's Committee. Melvene D. Hardee, Chairman. Cloth	4.50
Paper	3.75
Fifty-ninth Yearbook, 1960, Part I—*Rethinking Science Education*. Prepared by the Society's Committee. J. Darrell Barnard, Chairman. Cloth	4.50
Fifty-ninth Yearbook, 1960, Part II—*The Dynamics of Instructional Groups*. Prepared by the Society's Committee. Gale E. Jensen, Chairman. Cloth	4.50
Paper	3.75
Sixtieth Yearbook, 1961, Part I—*Development in and through Reading*. Prepared by the Society's Committee. Paul A. Witty, Chairman, Cloth	5.00
Paper	4.25
Sixtieth Yearbook, 1961, Part II—*Social Forces Influencing American Education*. Prepared by the Society's Committee. Ralph W. Tyler, Chairman. Cloth	4.50
Sixty-first Yearbook, 1962, Part I—*Individualizing Instruction*. Prepared by the Society's Committee. Fred T. Tyler, Chairman. Cloth	4.50
Sixty-first Yearbook, 1962, Part II—*Education for the Professions*. Prepared by the Society's Committee. G. Lester Anderson, Chairman. Cloth	4.50
Sixty-second Yearbook, 1963, Part I—*Child Psychology*. Prepared by the Society's Committee. Harold W. Stevenson, Editor. Cloth	6.50
Sixty-second Yearbook, 1963, Part II—*The Impact and Improvement of School Testing Programs*. Prepared by the Society's Committee. Warren G. Findley, Editor. Cloth	4.50
Sixty-third Yearbook, 1964, Part I—*Theories of Learning and Instruction*. Prepared by the Society's Committee. Ernest R. Hilgard, Editor. Cloth	5.50
Sixty-third Yearbook, 1964, Part II—*Behavioral Science and Educational Administration*. Prepared by the Society's Committee. Daniel E. Griffiths, Editor. Cloth	4.50
Sixty-fourth Yearbook, 1965, Part I—*Vocational Education*. Prepared by the Society's Committee. Melvin L. Barlow, Editor. Cloth	5.00
Sixty-fourth Yearbook, 1965, Part II—*Art Education*. Prepared by the Society's Committee. W. Reid Hastie, Editor. Cloth	5.00
Sixty-fifth Yearbook, 1966, Part I—*Social Deviancy among Youth*. Prepared by the Society's Committee. William W. Wattenberg, Editor. Cloth	5.50
Sixty-fifth Yearbook, 1966, Part II—*The Changing American School*. Prepared by the Society's Committee. John I. Goodlad, Editor. Cloth	5.00
Sixty-sixth Yearbook, 1967, Part I—*The Educationally Retarded and Disadvantaged*. Prepared by the Society's Committee. Paul A. Witty, Editor. Cloth	5.50
Sixty-sixth Yearbook, 1967, Part II—*Programed Instruction*. Prepared by the Society's Committee. Phil C. Lange, Editor. Cloth	5.00
Sixty-seventh Yearbook, 1968, Part I—*Metropolitanism: Its Challenge to Education*. Prepared by the Society's Committee. Robert J. Havighurst, Editor. Cloth	5.50
Sixty-seventh Yearbook, 1968, Part II—*Innovation and Change in Reading Instruction*. Prepared by the Society's Committee. Helen M. Robinson, Editor. Cloth	5.50
Sixty-eighth Yearbook, 1969, Part I—*The United States and International Education*. Prepared by the Society's Committee. Harold G. Shane, Editor. Cloth	5.50

	POSTPAID PRICE
Sixty-eighth Yearbook, 1969, Part II—*Educational Evaluation: New Roles, New Means.* Prepared by the Society's Committee. Ralph W. Tyler, Editor. Cloth....	$5.50
Sixty-ninth Yearbook, 1970, Part I—*Mathematics Education.* Prepared by the Society's Committee. Edward G. Begle, Editor. Cloth..............................	7.00
Sixty-ninth Yearbook, 1970, Part II—*Linguisitics in School Programs.* Prepared by the Society's Committee. Albert H. Marckwardt, Editor. Cloth................	5.50
Seventieth Yearbook, 1971, Part I—*The Curriculum: Retrospect and Prospect.* Prepared by the Society's Committee. Robert M. McClure, Editor. Cloth..........	5.50
Seventieth Yearbook, 1971, Part II—*Leaders in American Education.* Prepared by the Society's Committee. Robert J. Havighurst, Editor. Cloth....................	6.50
Seventy-first Yearbook, 1972, Part I—*Philosophical Redirection of Educational Research.* Prepared by the Society's Committee. Lawrence G. Thomas, Editor. Cloth.	5.50
Seventy-first Yearbook, 1972, Part II—*Early Childhood Education.* Prepared by the Society's Committee. Ira J. Gordon, Editor. Cloth........................	6.50

Distributed by
THE UNIVERSITY OF CHICAGO PRESS, CHICAGO, ILLINOIS 60637
1972

Please direct all orders for books to the University of Chicago Press.

S.C.E.R.T., West Bengal
Library
Calcutta
25/3, B. C. Road